In silico characterization of a rare disease that affects the dynamics and function of linker histone H1

Serhan TURUNÇ[1,2], Merve UÇA[1,2], Stefan DIMITROV[1,3], Dimitar ANGELOV[1,4], Seyit KALE[1,*]
[1]Izmir Biomedicine and Genome Center, Dokuz Eylül University Health Campus, Balçova, Izmir, Turkey
[2]Izmir International Biomedicine and Genome Institute, Dokuz Eylül University Health Campus, Balçova, Izmir, Turkey
[3]Université Grenoble Alpes, CNRS UMR 5309, INSERM U1209, Institute for Advanced Biosciences (IAB), Site Santé – Allée des Alpes, 38700, La Tronche, France
[4]Université de Lyon, Ecole Normale Supérieure de Lyon, CNRS, Laboratoire de Biologie et de Modélisation de la Cellule LBMC, 46 Allée d'Italie, 69007, Lyon, France
*Correspondence: seyit.kale@ibg.edu.tr

The eukaryotic genome is wrapped in nucleosomes to form chromatin. Nucleosomes contain 147 bp of DNA wrapped ~1.65 times around a histone octamer comprising two copies each of the core histones H2A, H2B, H3, and H4. In higher eukaryotes, a fifth histone, known as the linker histone H1, binds the nucleosome core particle and the linker DNA between nucleosomes so as to facilitate the delicate folding of the chromatin fiber into higher order three-dimensional structure. H1 is a tripartite protein composed of a folded central region sandwiched between a short N-terminal domain and a long, highly positively charged C-terminal domain (CTD). The long unstructured carboxy-terminal domain (CTD) of linker histone is poorly conserved among subtypes and species. In mammals, 11 H1 variants have been identified, including seven somatic H1s and four germ-cell-specific H1s. Linker histone H1.4, which belongs to the somatic subtypes of family H1, plays a crucial role in the maintenance of the stable higher-order structure of chromatin.

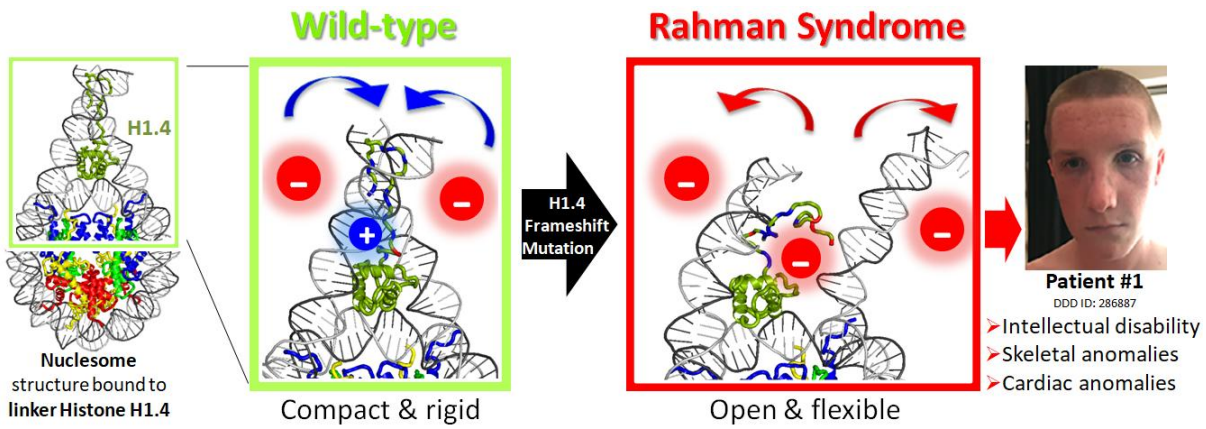
In late 2019, germline frameshift mutations involving the CTD of H1.4, were causally linked to an as-yet poorly defined syndrome, later called the Rahman syndrome (RS). In affected individuals, characteristic phenotypes include intellectual disability, skeletal and cardiac anomalies. All disease-causing mutations are located in a small region in the CTD tail and generally lead to a shorter protein with a highly negatively charged C-terminal tail. Motivated by this fact, we hypothesize that the chromatin domains harboring the RS H1.4 mutant could exhibit a less condensed and differently organized structure, which could disrupt the delicate architecture and function of chromatin.

To understand how the Rahman syndrome (RS) mutant Histone H1.4 affects nucleosome dynamics, we performed atomistic molecular simulations of the WT (wild type) - as well as the RS mutant H1.4 bound nucleosome core particles at the sub-microsecond timescale. Because the complete CTD is too large for atomistic-level simulations, we incorporated a ~30 amino acid segment close to the globular domain known to be critical for the compaction function of H1. Among all RS patients, we chose the very first patient, "patient #1", because of the early positioning of the frameshift mutation in the genotype and the dramatic net charge reversal of the CTD. Indeed, we found that the H1.4 of patient #1 could not only not fold the linker DNA strands but it was actually destabilizing them in comparison to the WT as well as a "tailless" mutant that lacks the CTD altogether.

In all surviving RS patients, there is a small shared region at the terminus of the mutated CTD tail which is elusive in terms of its cytotoxic role as well as potential loss- or gain of function in the cell. To characterize this region in detail, we performed accelerated atomistic simulations of the isolated shared mutant CTD segment as well as its WT counterpart. In these simulations, also known as metadynamics, we encourage the disordered protein to visit rare and novel structural configurations more frequently than it would normally do under ambient equilibrium conditions. We found that the mutant segment is distinctly more compact in comparison to its WT counterpart, which could be potentially implicated in the binding of de novo intra-nuclear partners.

In conclusion, rare diseases are not rare in the increasingly crowded world population, especially for Turkey, as a Mediterranean country. The findings of this study will not only give us information about the Rahman syndrome but also provide important findings about the 3D chromatin architecture mechanism, which is still a mystery in the epigenetics field. In the near future, we will apply EuroHPC for extra computing time. It will accelerate the analysis of complete Rahman syndrome CTD simulations. In this project, *in cellulo* and *in vitro* experiments are still a work in progress in parallel with our *in silico* research. The combination of these studies is expected not only to shed light on the mechanism of Rahman syndrome but could also be used as an example of a "workflow" for rare disease studies.

Acknowledgement and/or disclaimers, if any:
This project is funded by the European Molecular Biology Organization (EMBO) Installation Grant No 5056, Tübitak-Bideb 2247-A project no: 120C149, Tübitak ULAKBIM Truba High Performance Computing resource, Partnership for Advanced Computing in Europe (PRACE) project DECI-17-EPICENTROMERE, as well as local high performance computing resources of Izmir Biomedicine and Genome Center.


Wild-type
Rahman Syndrome
H1.4
Nucleosome
structure bound to
linker Histone H1.4
Compact & rigid
H1.4
Frameshift
Mutation
Open & flexible
Patient #1
DDD ID: 286887
Intellectual disability
Skeletal anomalies
Cardiac anomalies