Assessing the limits of restraint-based 3D modeling of genomes and genomic domains

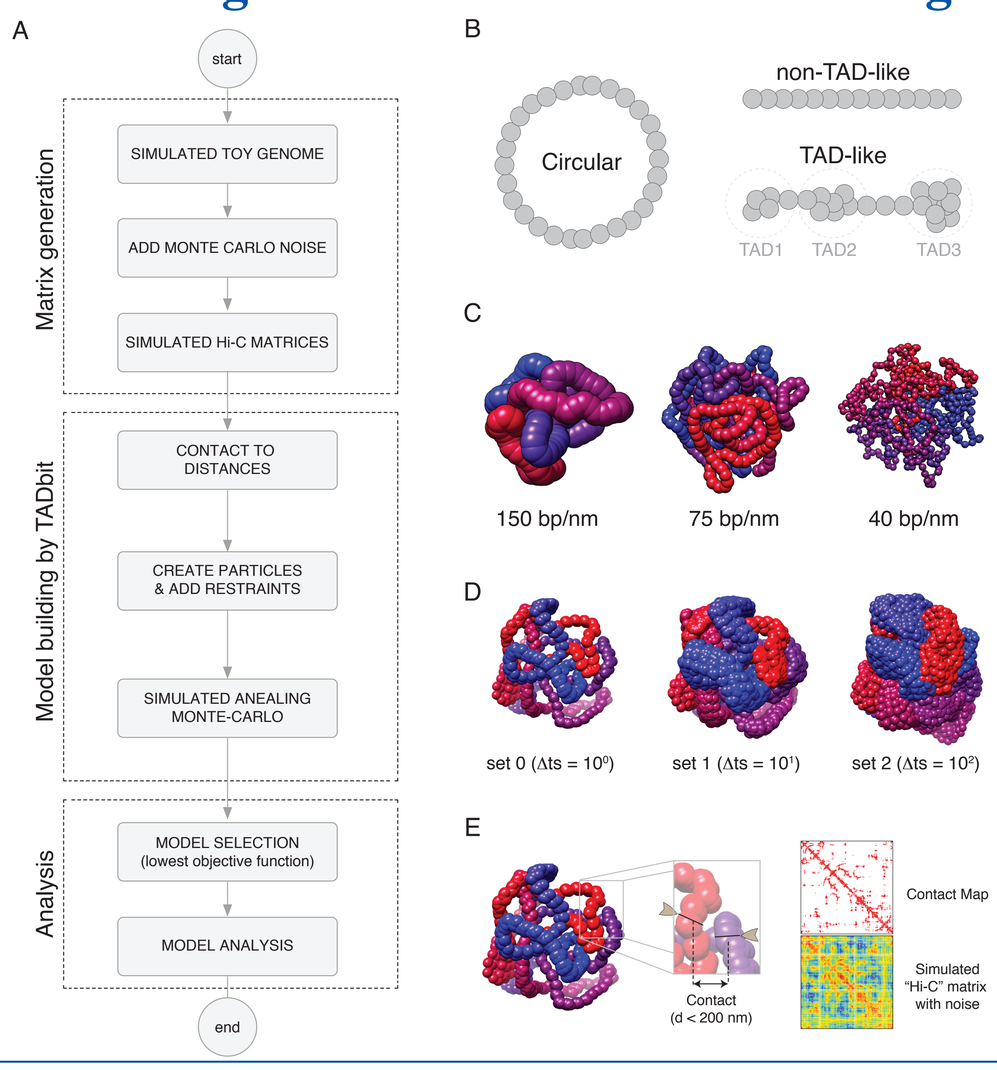
***Trussart M.*[1,2] , *Serra F.*[3,4], *Baù D.*[3] , *Junier I.*[2,4], *Serrano L.*[1,2,5], *Marti-Renom M.A.*[3,5*]**

1. EMBL/CRG Systems Biology Research Unit, Centre for Genomic Regulation (CRG), Barcelona, Spain ; 2. Universitat Pompeu Fabra (UPF), Barcelona, Spain ;

3. Structural Genomics Group. Centre Nacional d'Anàlisi Genòmica - Centre de Regulació Genòmica (CNAG-CRG). Barcelona, Spain. ; 4. Gene Regulation, Stem Cells and Cancer Program, Centre for Genomic Regulation (CRG), Barcelona, Spain ; 5. Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

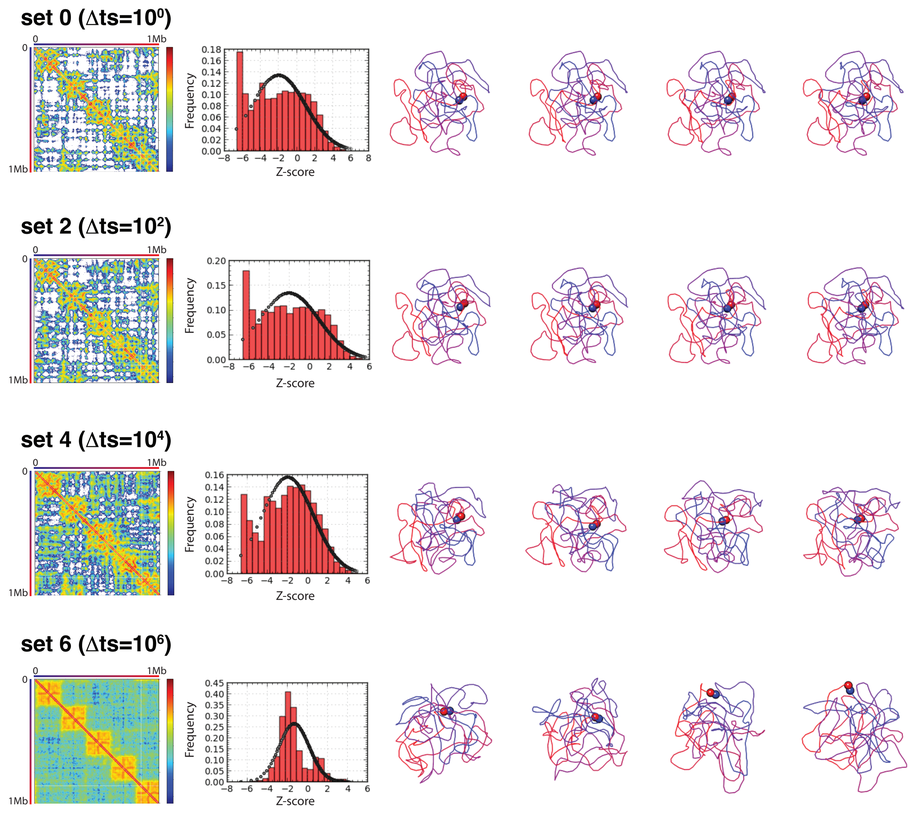
Restraint-based modeling of genomes has been recently explored with the advent of Chromosome Conformation Capture (3C-based) experiments. We previously developed a reconstruction method to resolve the 3D architecture of both prokaryotic and eukaryotic genomes using 3C-based data. These models were congruent with fluorescent imaging validation. However, the limits of such methods have not systematically been assessed. Here we propose the first evaluation of a mean field restraint-based reconstruction of genomes by considering diverse chromosome architectures and different levels of data noise and structural variability.

Matrix generation and model building



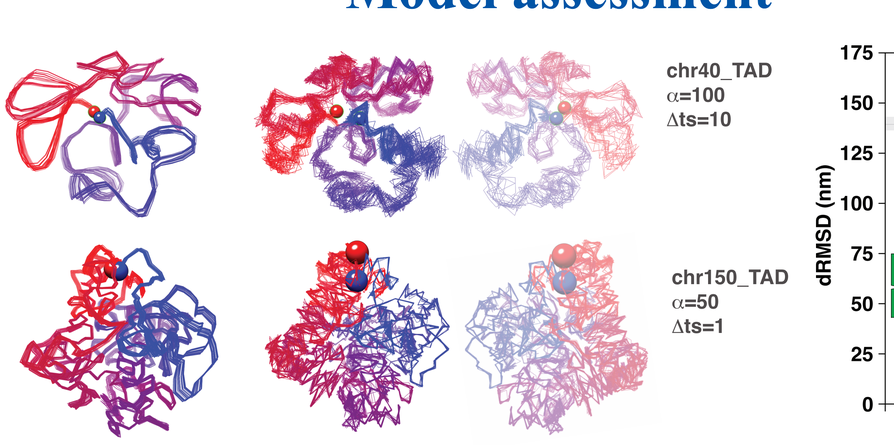
1. Using polymer modeling, we simulated "toy genomes" with different architectures, from which we extracted simulated interaction matrices with increasing noise levels and structural diversity.

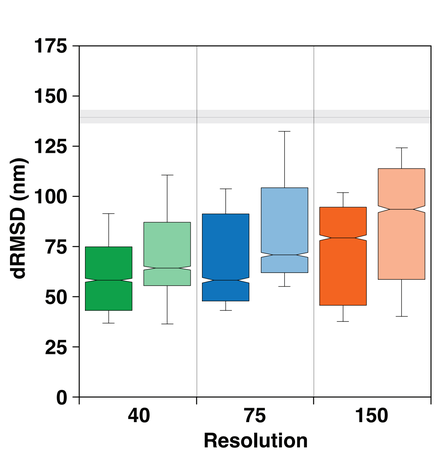
Simulated Hi-C interaction matrices

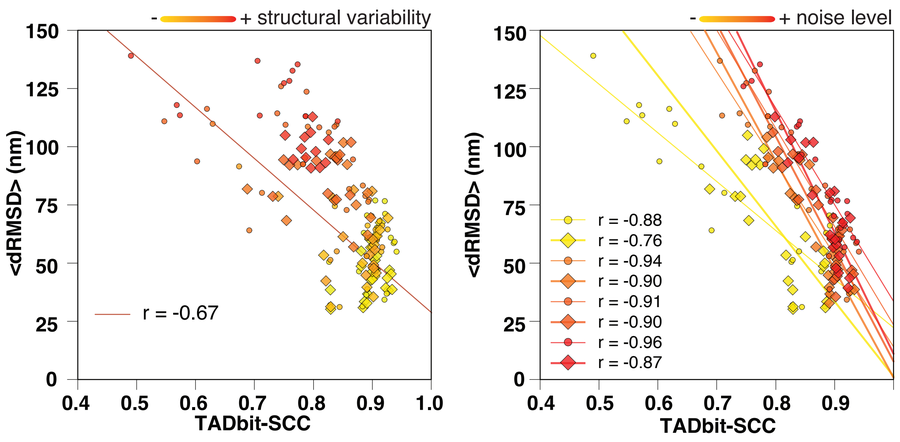


2. We reconstructed with TADbit 3D models based on the simulated "Hi-C" matrices. 3. We analyzed the reconstructed models to assess their structural similarity to the original simulated toy genomes.

Model assessment







4. The accuracy of the models can be *a priori* predicted by MMP score from the properties of the Hi-C matrix

Matrix Modeling Potential (MMP) score