

# A visual grammar for 3DGenomics

#### Marc A. Marti-Renom CNAG-CRG · ICREA

Unpublished







Clementina Altube ELISAVA







ELISAVA



Santiago Bonet Garrote ELISAVA

Francesc Ribot Puntí



Jonathan Chacón Pérez ELISAVA

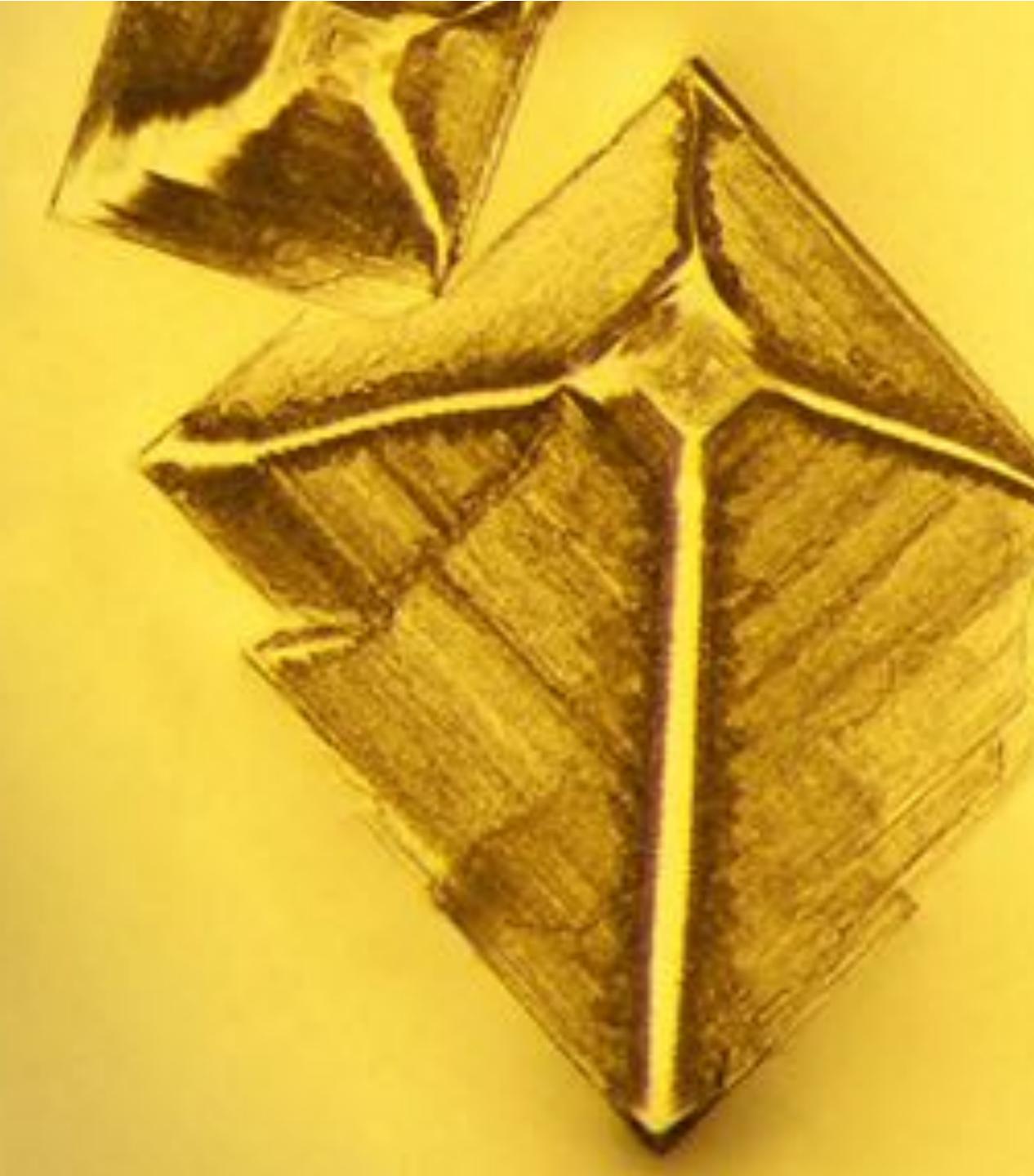
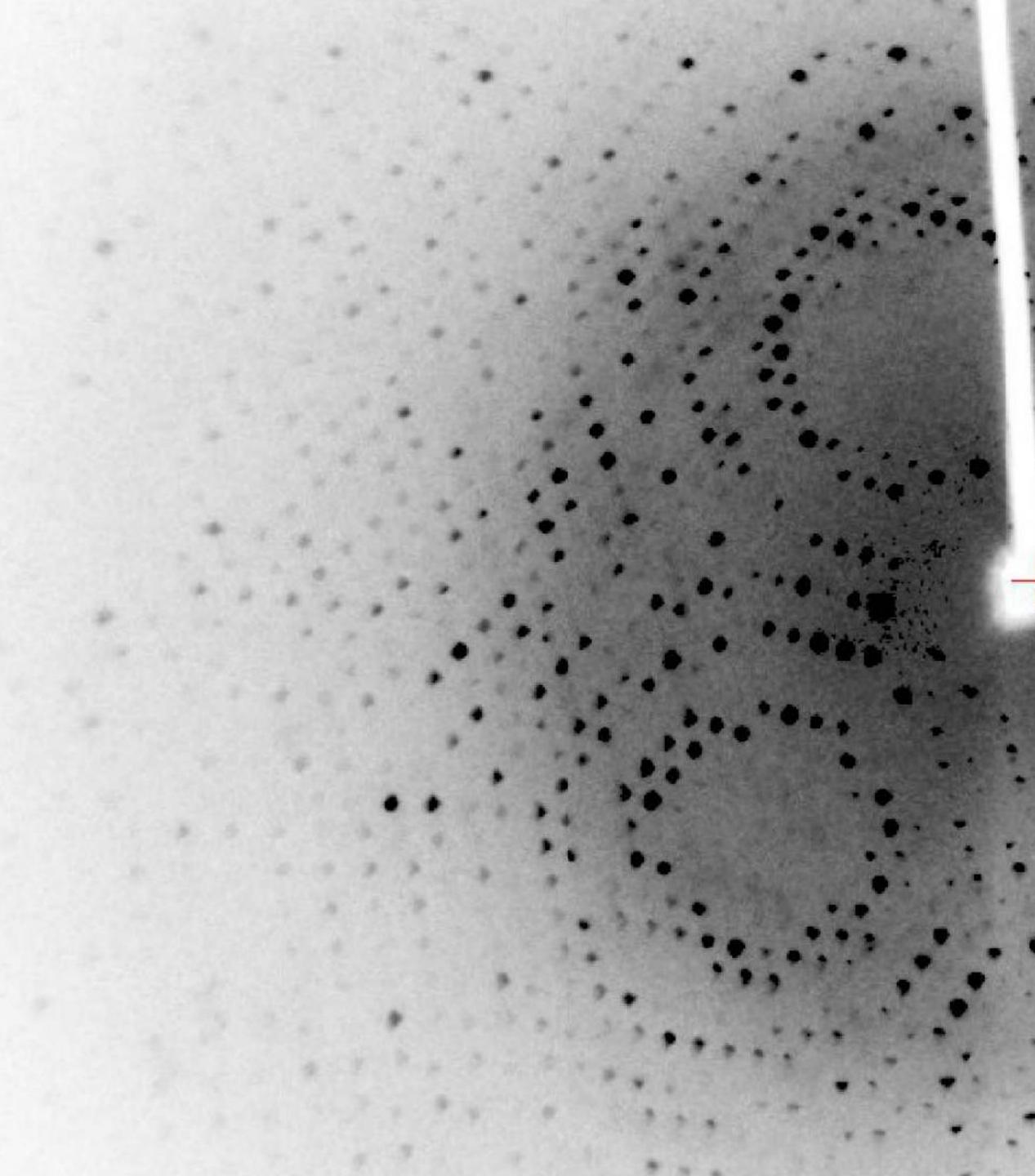
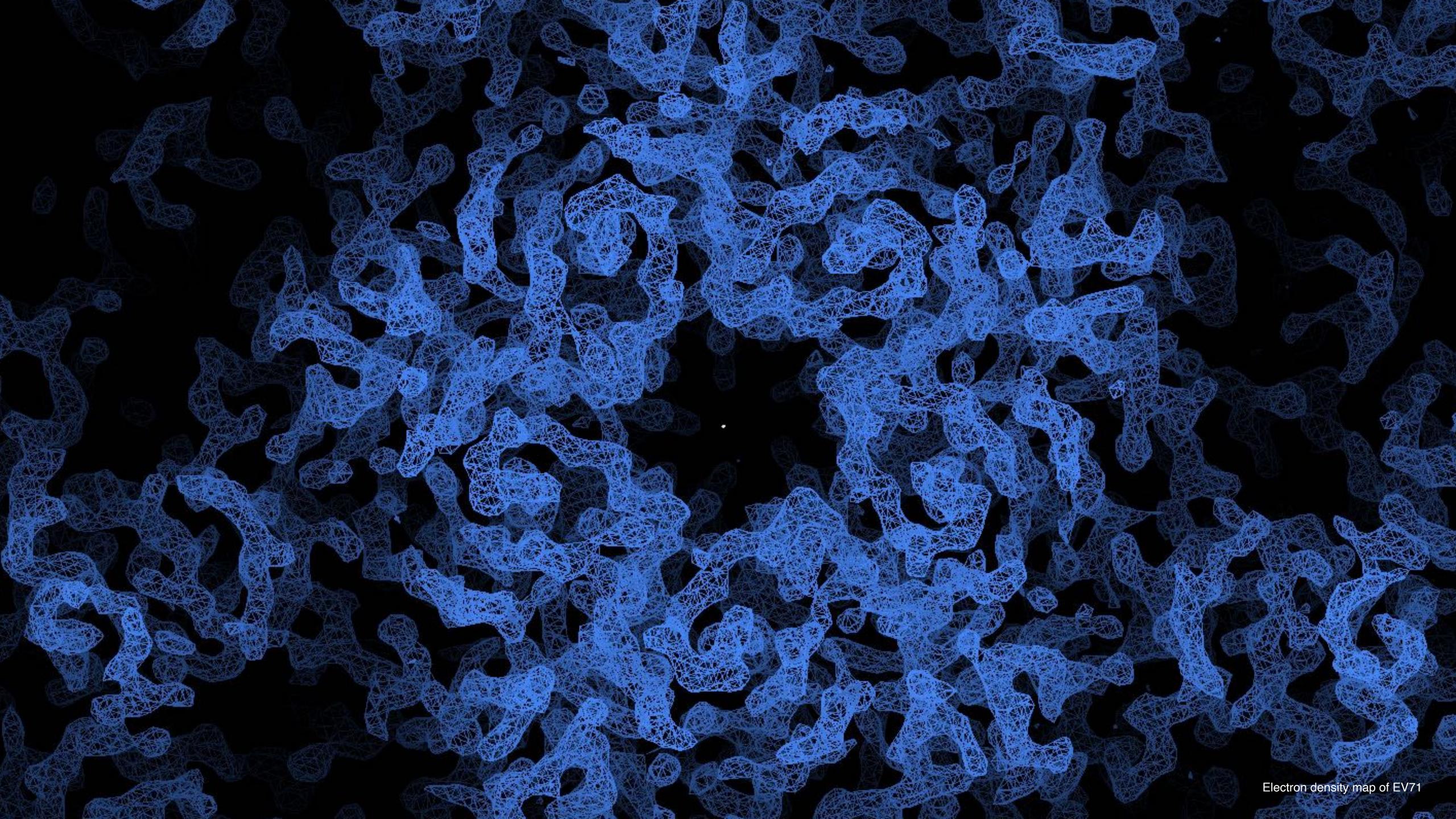


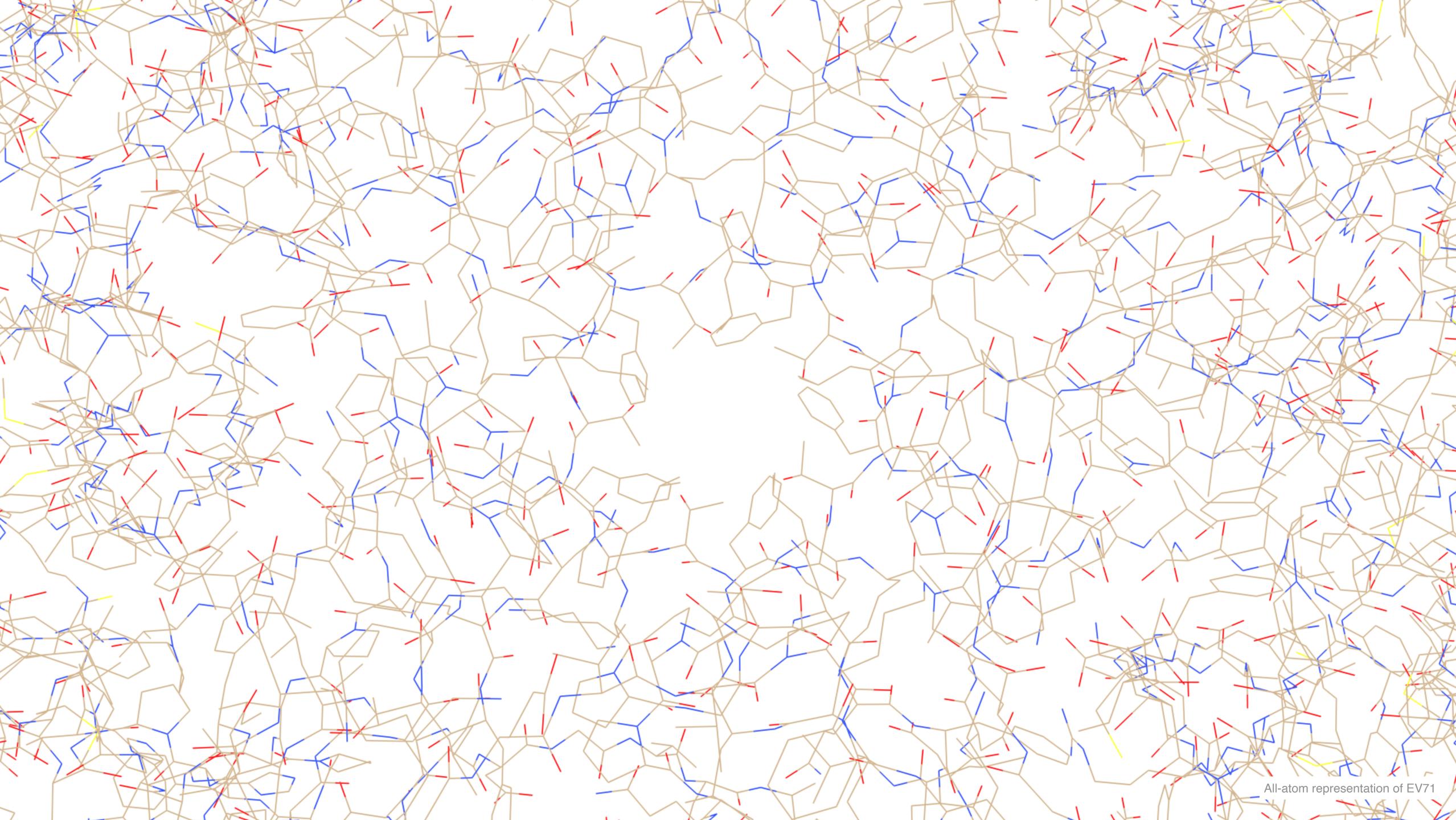
image credit https://ecm31.ecanews.org

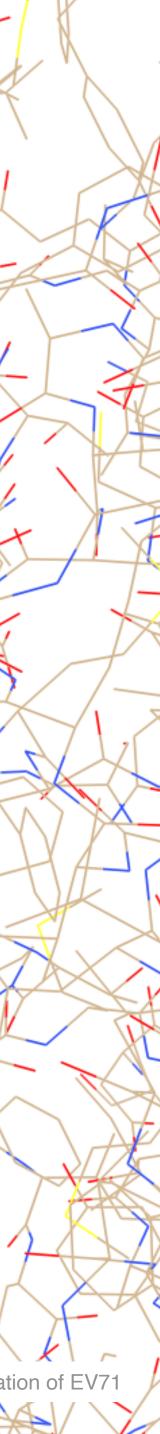


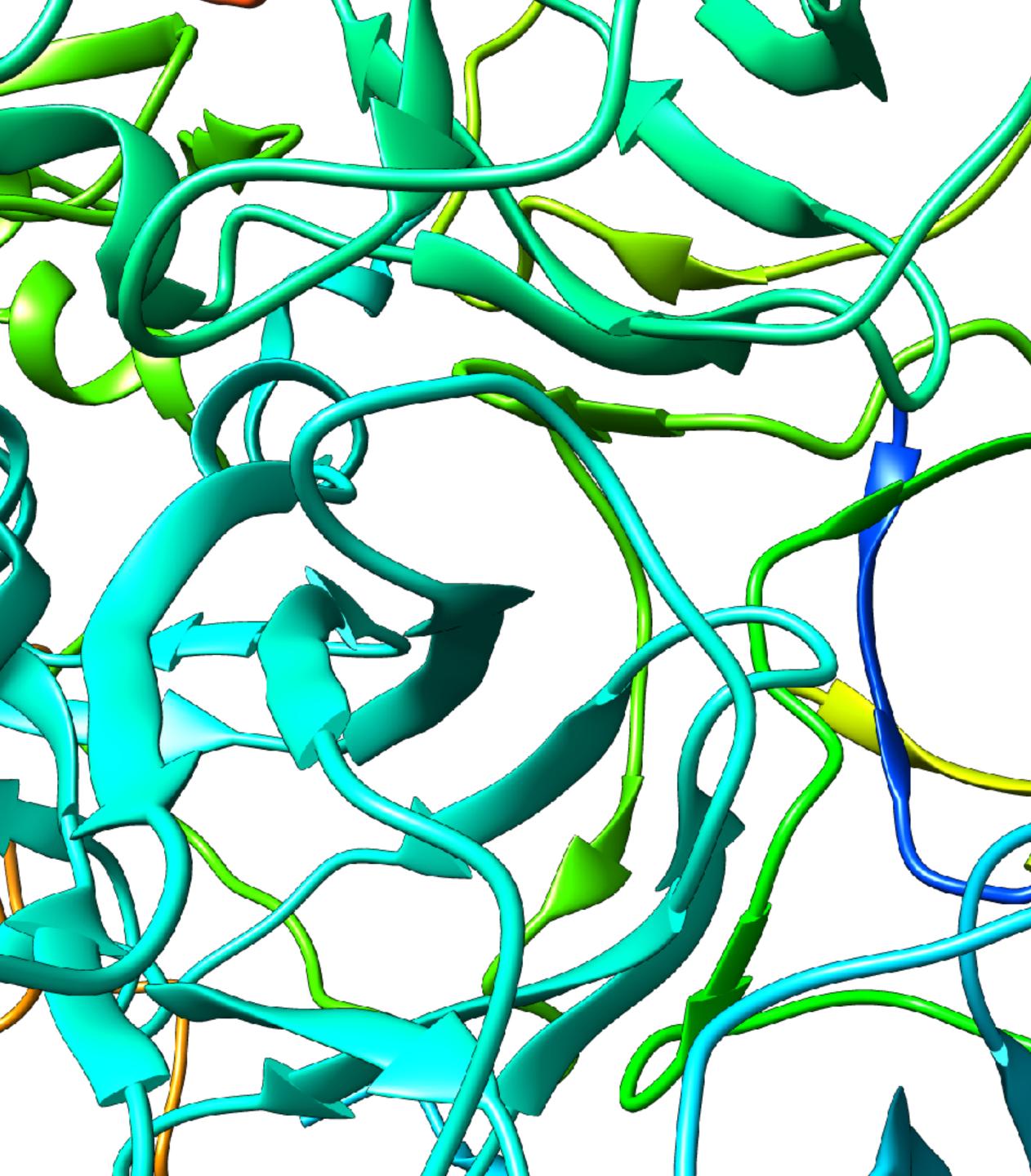


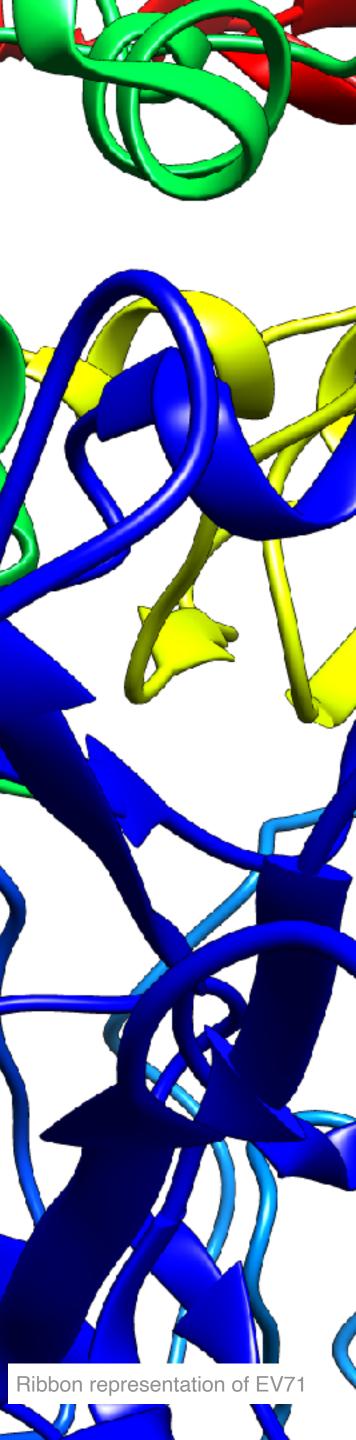


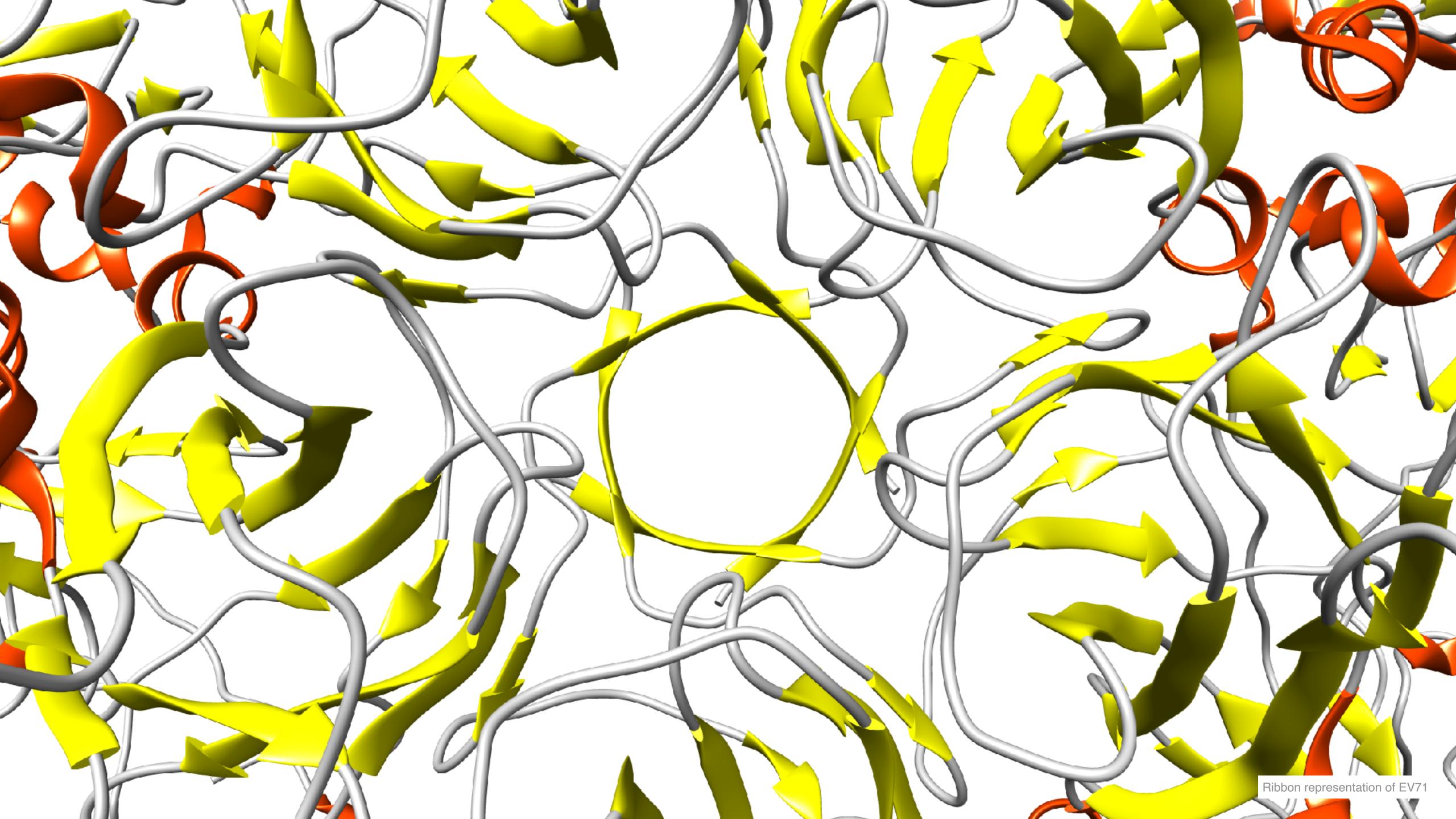


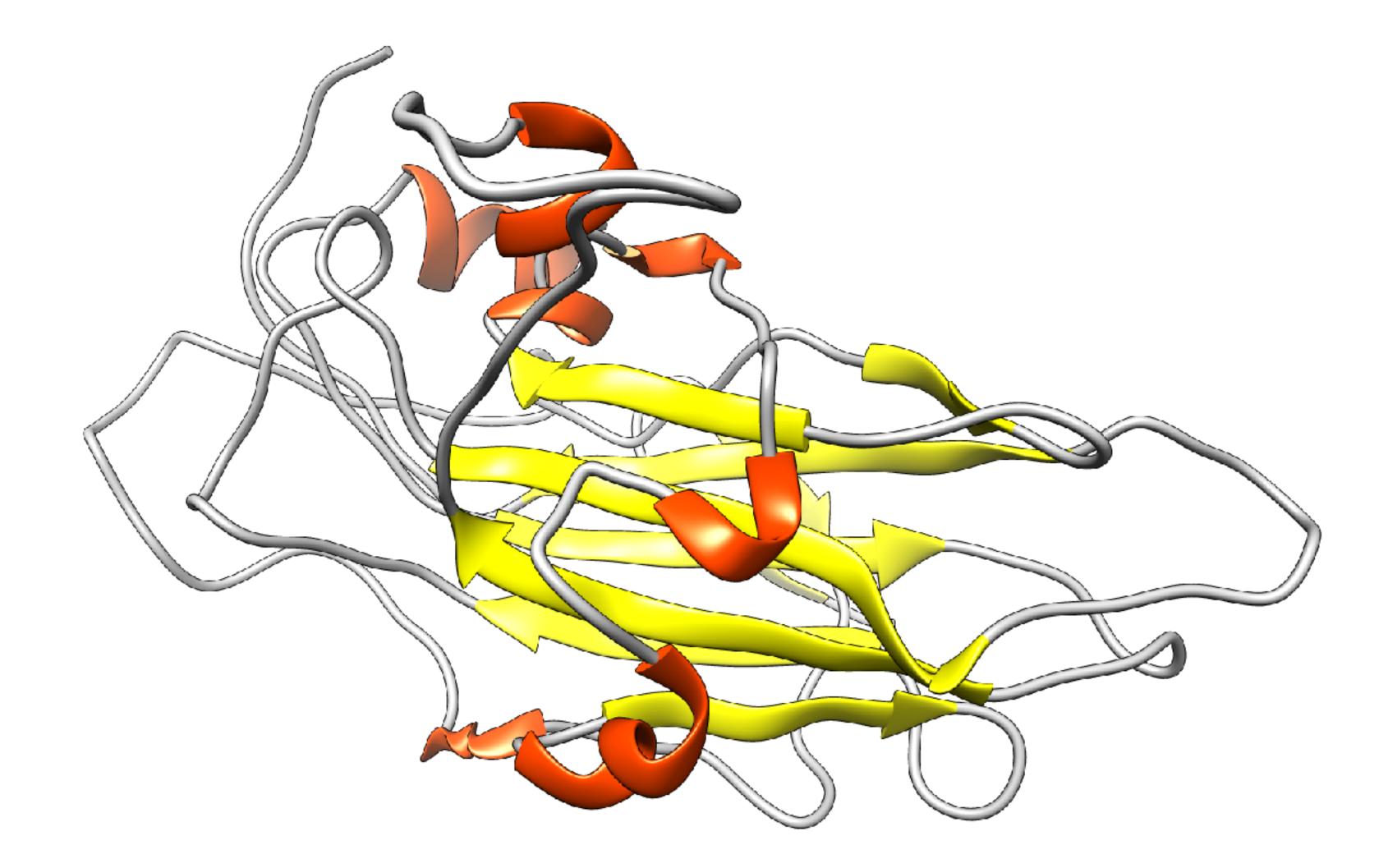




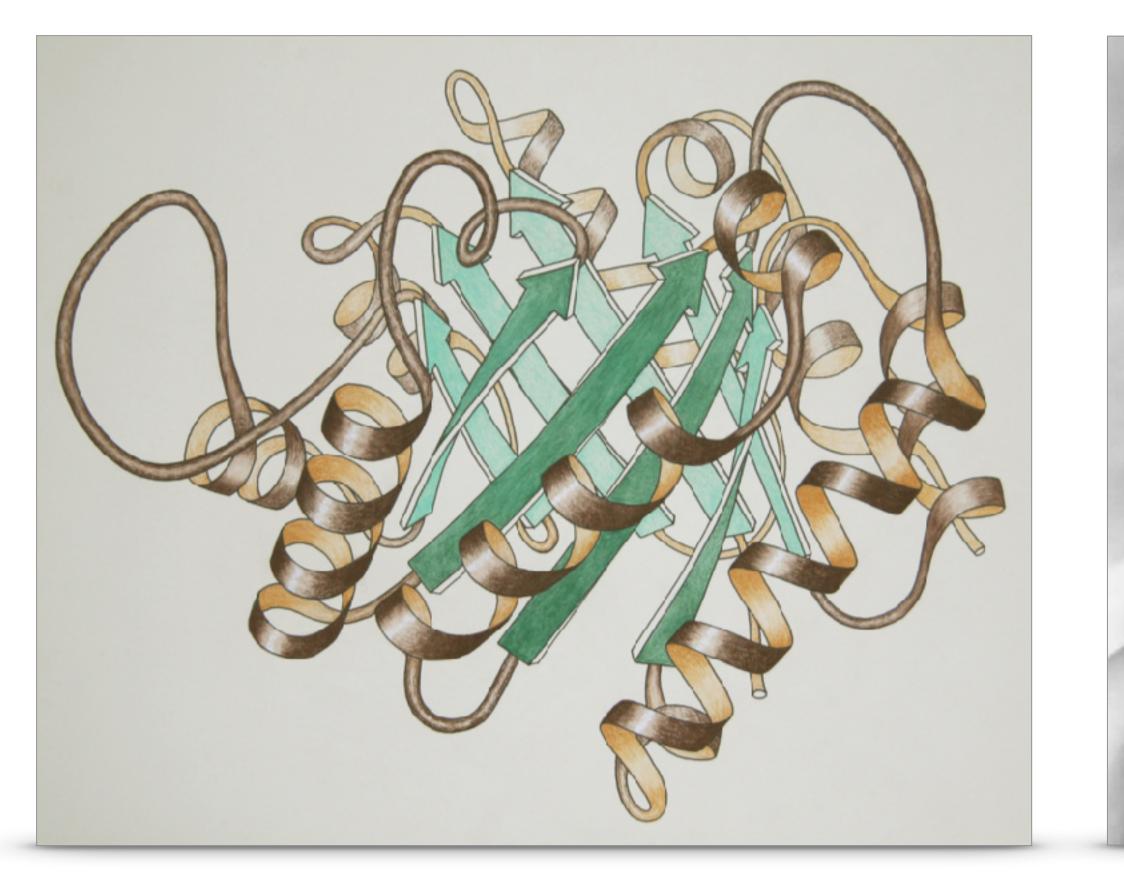








## Ribbon diagrams of proteins 70s-80s >> Dick Dickerson · Irving Geis · Bo Furugren · Jane S. Richardson





Solution at the second second

#### history

#### Early ribbon drawings of proteins

Jane S. Richardson

drawings' for our own staphylococcal contrast negatives for printing. nuclease and superoxide dismutase Making these drawings was a fascinat-structures, and for other proteins<sup>5</sup> ing process. First, the structures are very orientation clearer, smoothed spiral rib-(Fig. 1*a*). However, the first attempt to aesthetically pleasing — especially, for bons without thickness for helices, and illustrate the full range of known protein me, the varied and elegant curves of rounded 'ropes' for loops. Surprisingly, structures (only 75 different ones, then!)  $\beta$ -sheets. Second, making a drawing can these disparate parts look visually uniwith a consistent system of representa- change one's scientific understanding of fied and intelligible. Local orientation of tion was the 1981 article "The anatomy a protein, sometimes revealing a prefer- arrows was also smoothed in the direc-and taxonomy of protein structure" able structural classification and once tion between strands, to strengthen the (Fig. 1*b*–*d*).

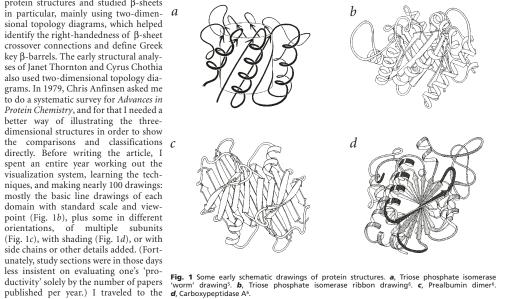
During the 1970s we had compared protein structures and studied  $\beta$ -sheets in particular, mainly using two-dimensional topology diagrams, which helped identify the right-handedness of  $\beta$ -sheet crossover connections and define Greek key  $\beta\text{-barrels}.$  The early structural analy ses of Janet Thornton and Cyrus Chothia also used two-dimensional topology dia grams. In 1979, Chris Anfinsen asked me to do a systematic survey for Advances in Protein Chemistry, and for that I needed a better way of illustrating the threedimensional structures in order to show the comparisons and classifications directly. Before writing the article, I spent an entire year working out the visualization system, learning the techniques, and making nearly 100 drawings: mostly the basic line drawings of each domain with standard scale and view point (Fig. 1*b*), plus some in different orientations, of multiple subunits (Fig. 1*c*), with shading (Fig. 1*d*), or with side chains or other details added. (Fortunately, study sections were in those days

The ubiquitous ribbon drawings of pro-tein structures that are commonly made and-white molecular display to choose from one residue to the next, which is the these days using programs such as viewpoint and print out  $C\alpha$  traces at a fundamental basis of 'ribbons', is context Molscript or Ribbons had their origin consistent scale. The drawings were made dependent. In a helix, the direction of ~20 years ago in drawings made by in pencil on tracing paper over the print- consecutive CO vectors is nearly parallel, hand. Some earlier schematic drawings out, while also looking at a smaller ver- in a  $\beta$ -strand it is nearly antiparallel, and had been made of individual proteins: sion of the structure in stereo view, and in loops it is often near 90° but has no Dick Dickerson was the first to make a finally traced in India ink. Shaded black- sensible structural or visual meaning. protein schematic1 and Irving Geis the and-white (Fig. 1d) or color versions Several later computer-based system first to show successive peptide planes were made by sticking on pieces of over- tried using a consistent definition for all with ribbons<sup>2</sup>; ribbon drawings, using lay film and cutting to fit the ribbon parts, but they were soon abandoned various conventions, were made by a few edges. Dave monitored the blackness of because the results looked confusing (for others<sup>3</sup>, most notably Bo Furugren<sup>4</sup>. My my ink, touched up line quality under a example,  $\beta$ -strands flipping over every husband David and I had made 'worm microscope, and photographed high-residue).

tion was surprisingly complex and interesting (for more on this aspect, see ref. 8). Not only were those conventions modified from various precursors and elaborated in new ways, but there is an inherent logical conflict that dictates a ertain level of inco Specifically, the definition for how to

even correcting a chain tracing<sup>7</sup>. Third, defining the conventions of representa-

I finally chose smoothed arrows for

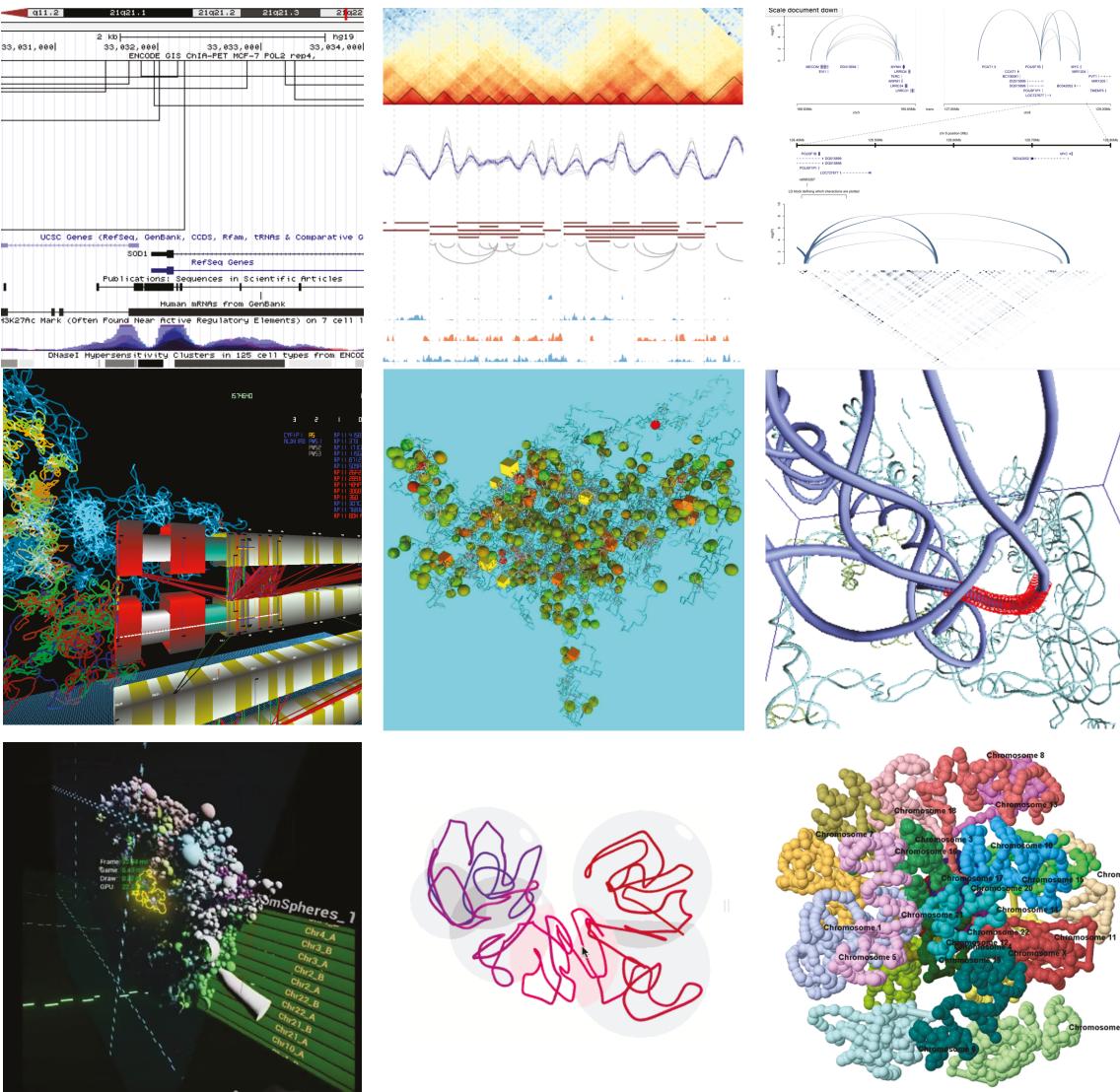


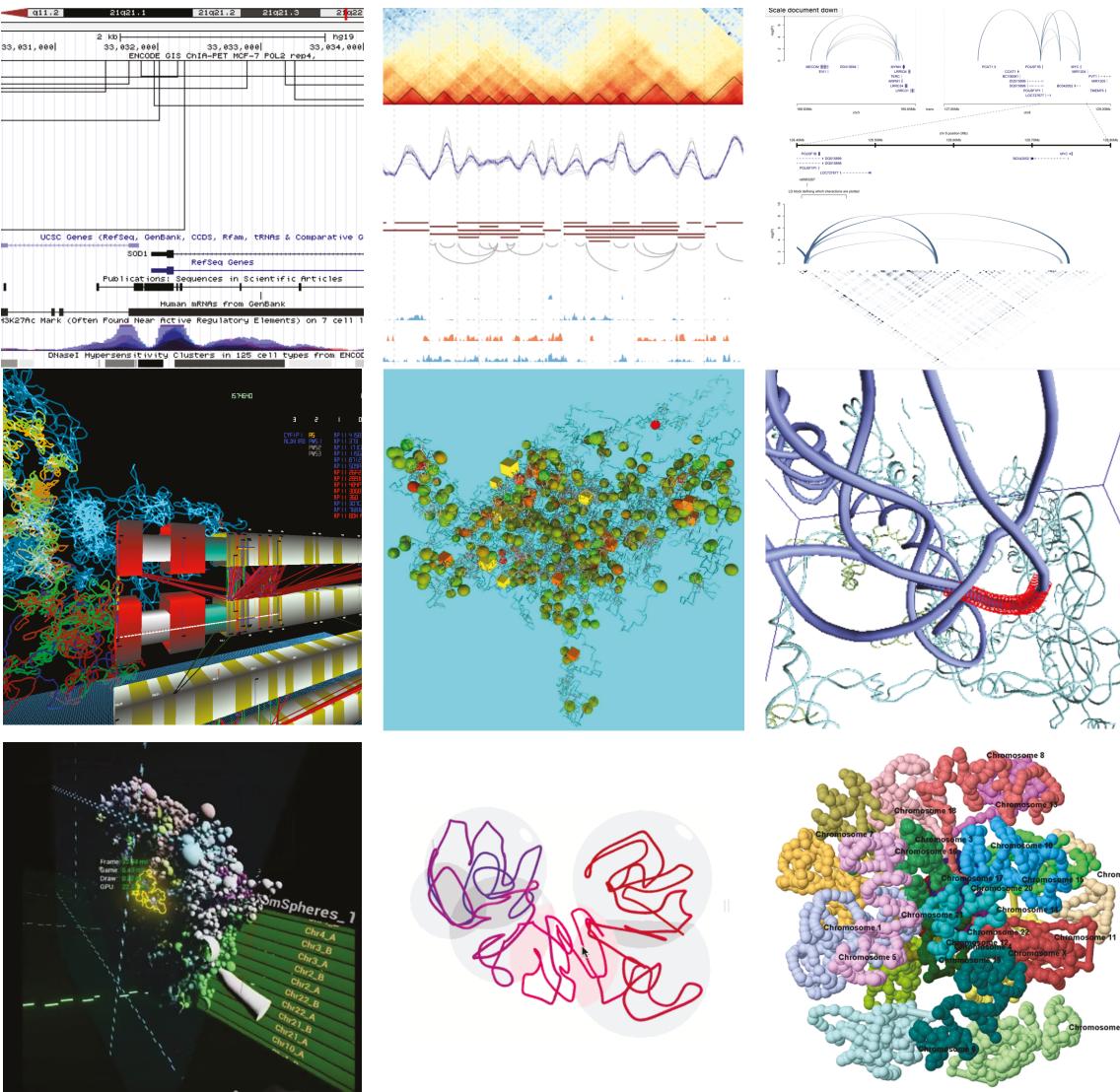
nature structural biology • volume 7 number 8 • august 2000

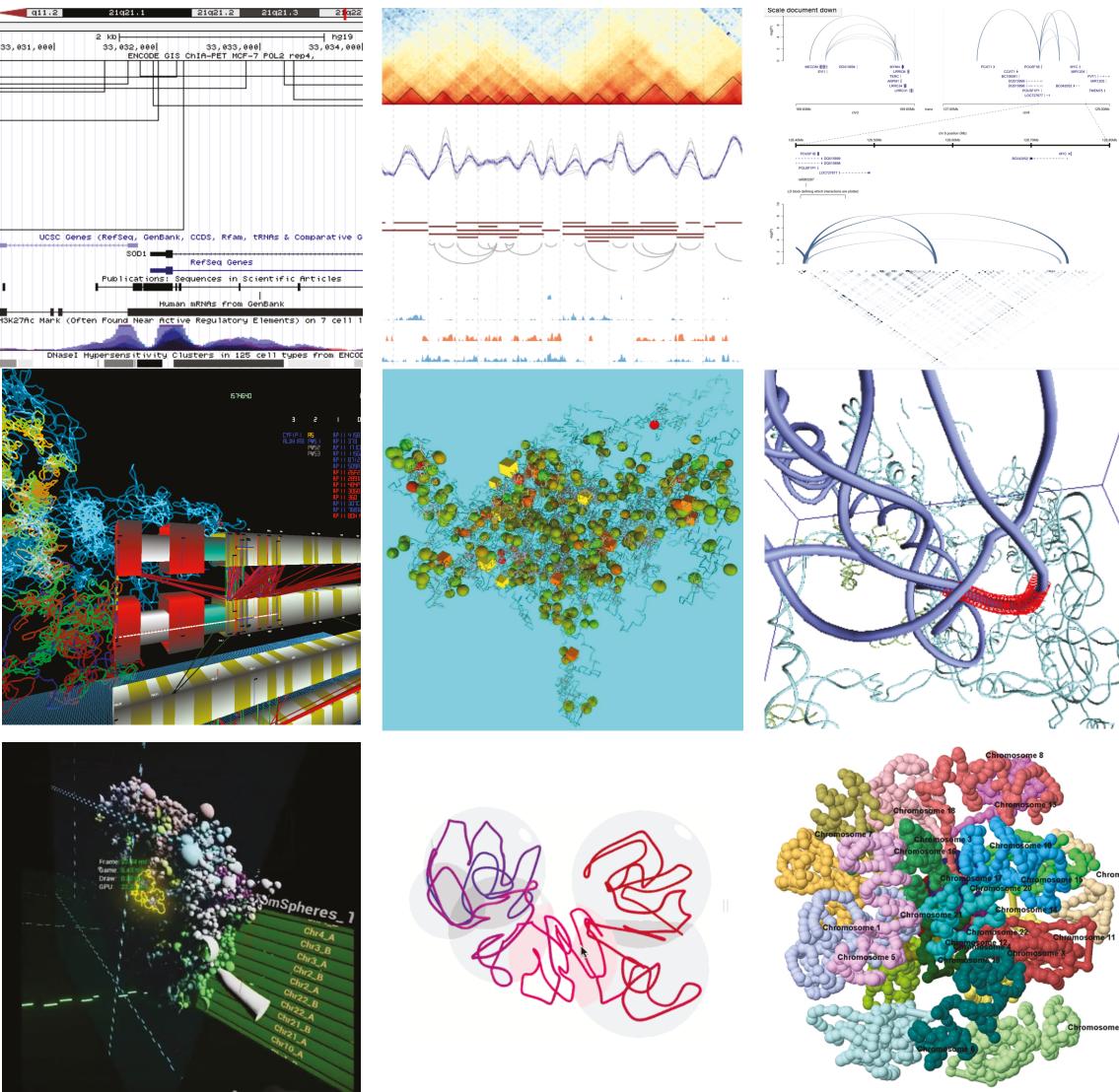
Ribbon schematic of triose P isomerase monomer (hand-drawn by J. Richardson, 1981) Richardson, Jane S. (1981), "Anatomy and Taxonomy of Protein Structures", Advances in Protein Chemistry, Advances in Protein Chemistry, 34: 167–339 Richardson, Jane S. (2000), "Early ribbon drawings of proteins", Nature Structural Biology, 7 (8): 624–625,

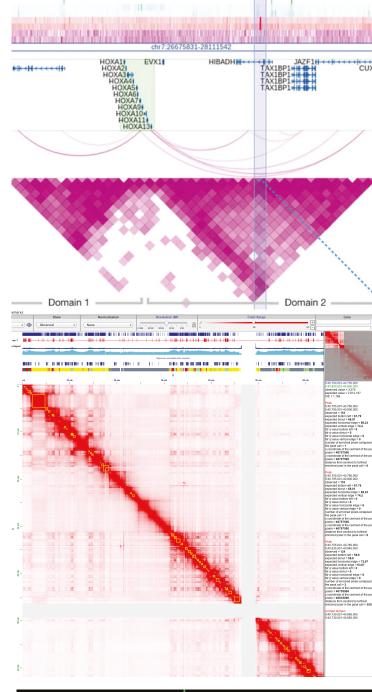


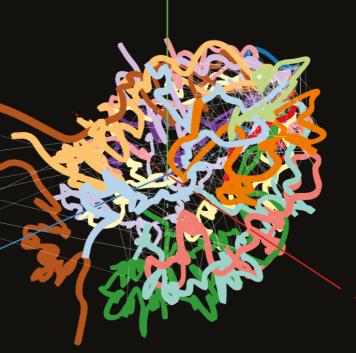
#### 3D Genome orgainzation how we represent it...



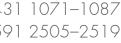




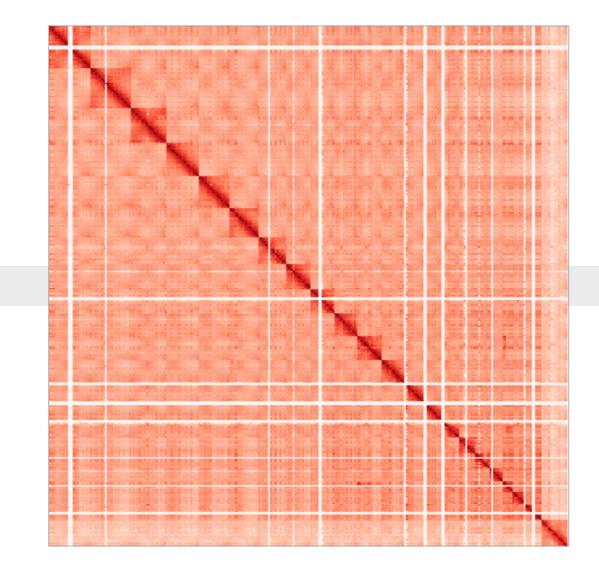


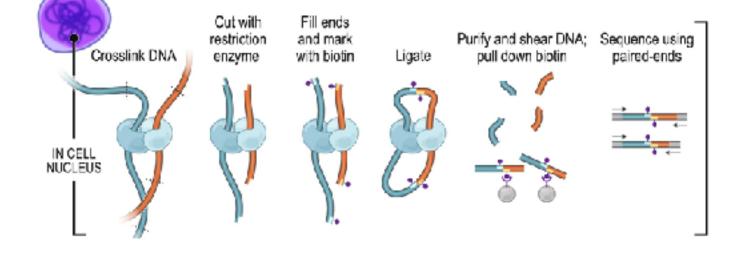


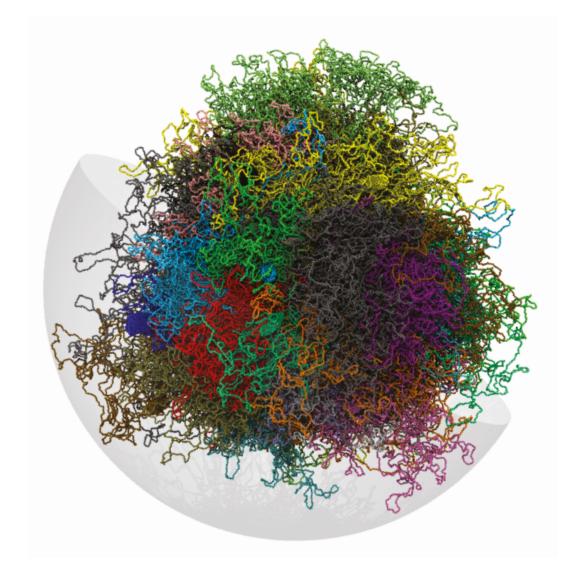
Goodstadt, M. and Marti-Renom, M.A. "Biovisualization of the Genome, from Data Analysis and Hypothesis Generation to Communication and Learning" Journal of Molecular Biology (2018) 431 1071–1087 Goodstadt, M. and Marti-Renom, M.A. "Challenges for visualizing three-dimensional data in genomic browsers" FEBS Letters (2017) 591 2505–2519



## 3D Genome orgainzation what we know...



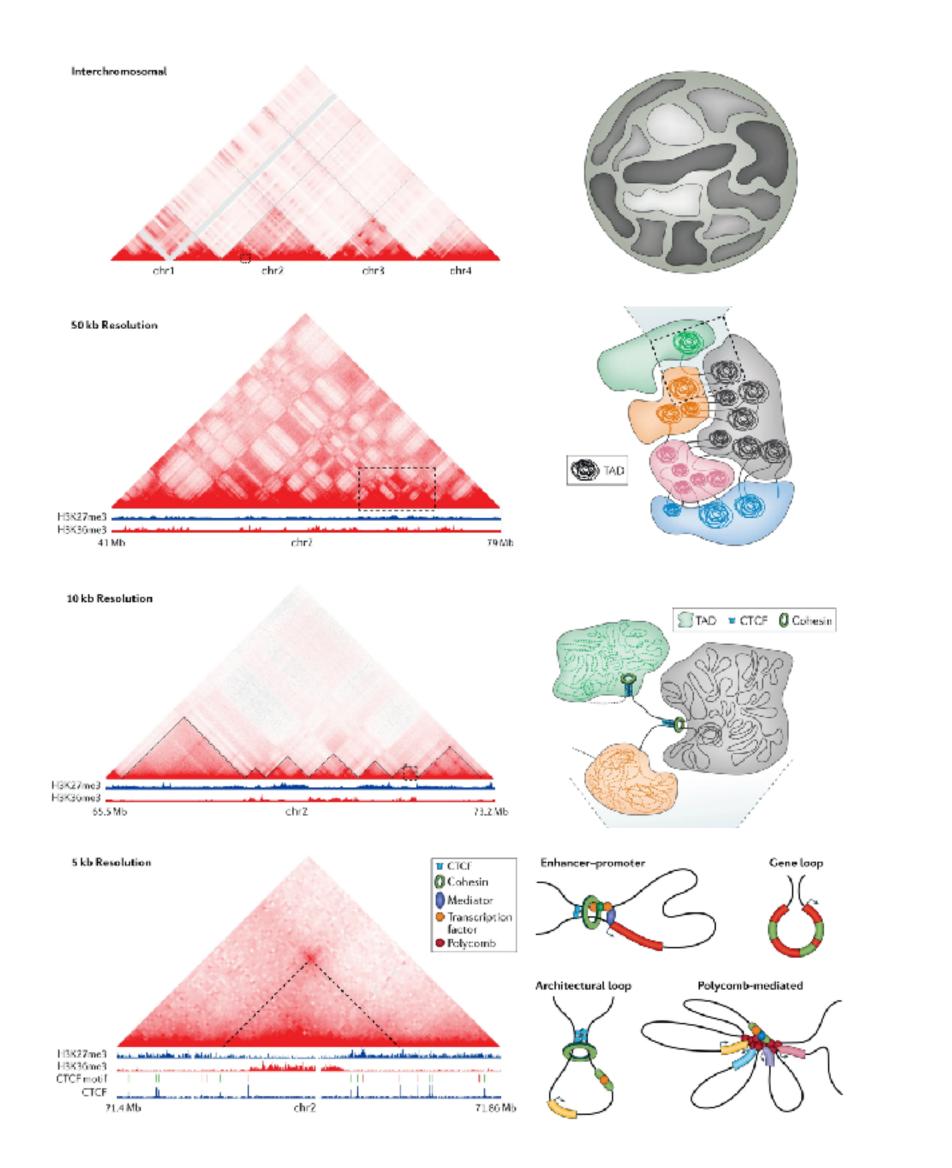


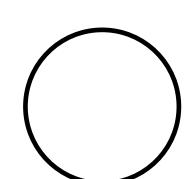


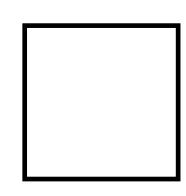
Lieberman-Aiden et al. Science 2009 Oct 9;326(5950):289-93 Baù & Marti-Renom Chromosome Res. 2011 Jan; 19(1):25-35 Di Stefano et al. Scientific Reports 2016 Oct 27 (6):35985

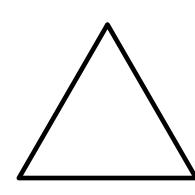


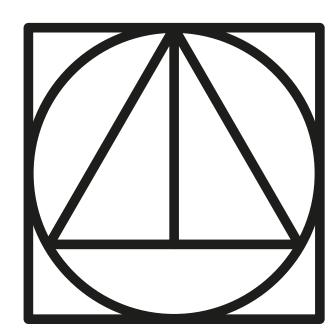
## 3D Genome orgainzation what we know...









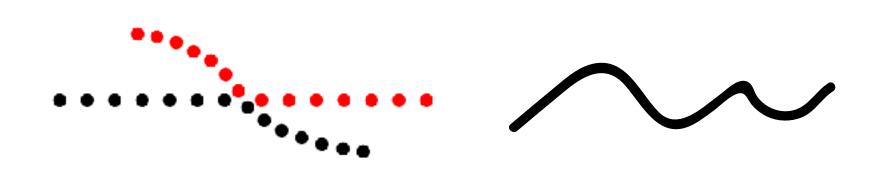


# Geometric Diagrams of Genomes

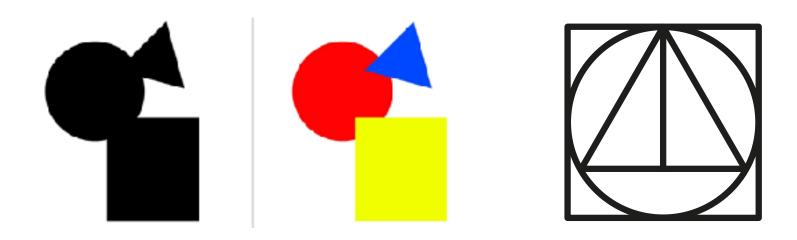
a visual grammar for 3DGenomics

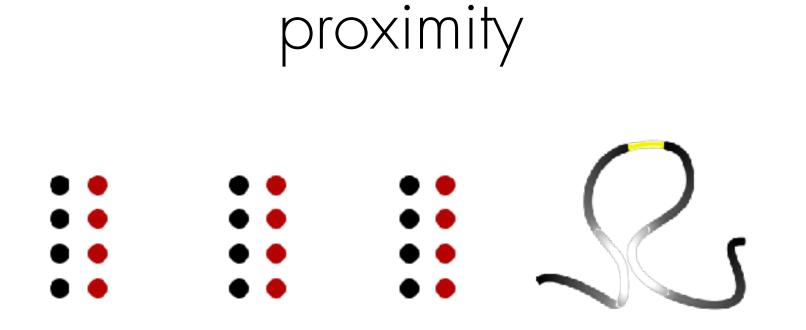
#### (some) The Gestalt Principles

#### continuation



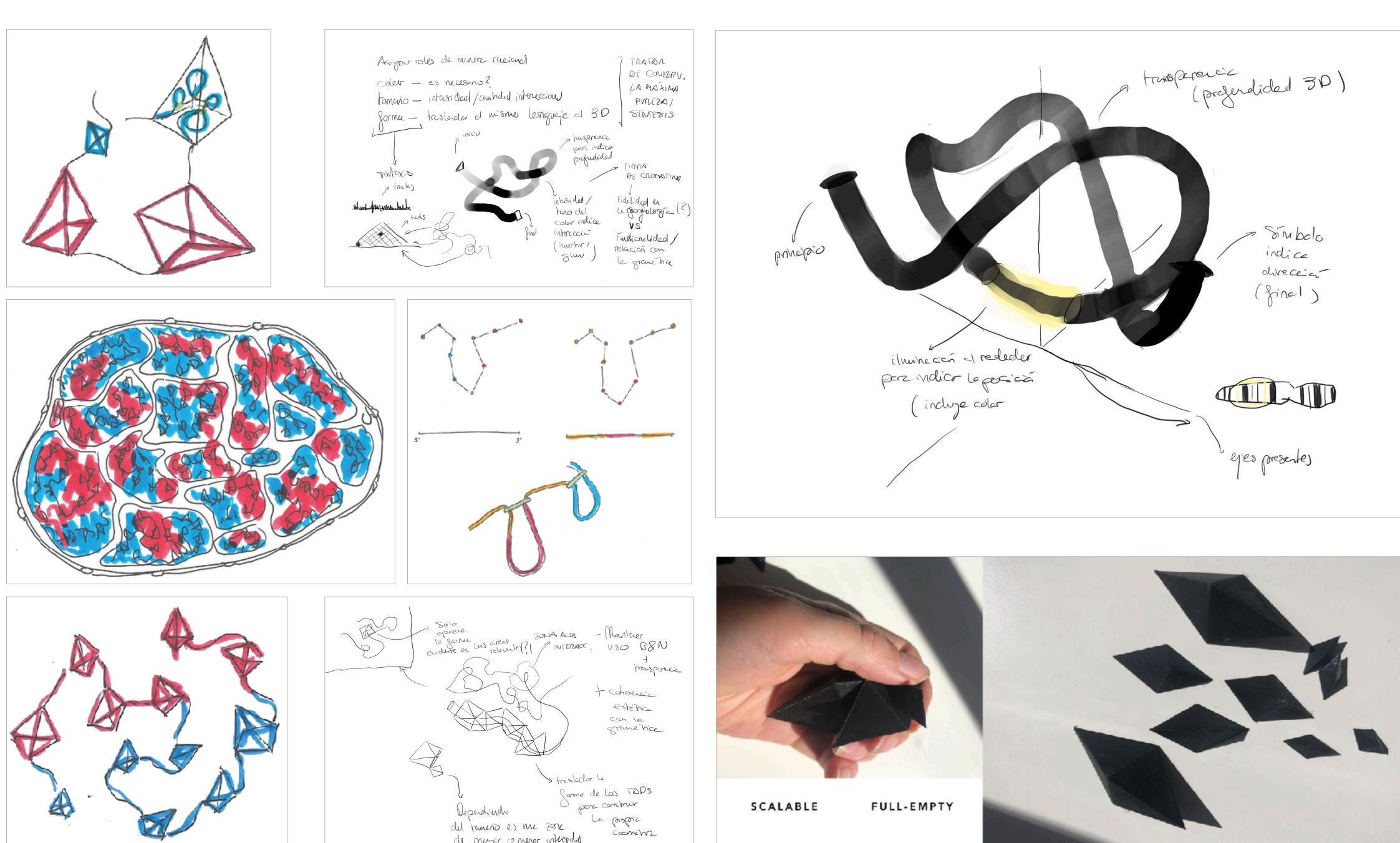
### symmetry and order

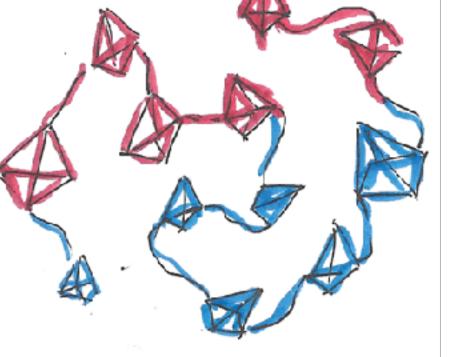




## focal point

## 



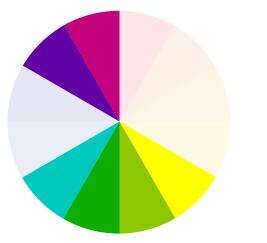




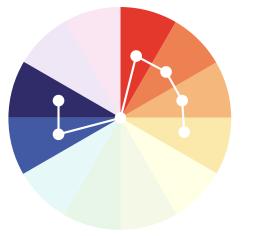


# Visual Grammar for 3DGenomics color

high-intensity (bright colors)



cold to warm (muted colors)



silent







structure (3D) 3C-based data · 3D models

## sequence (1D)

genes  $\cdot$  genome annotation

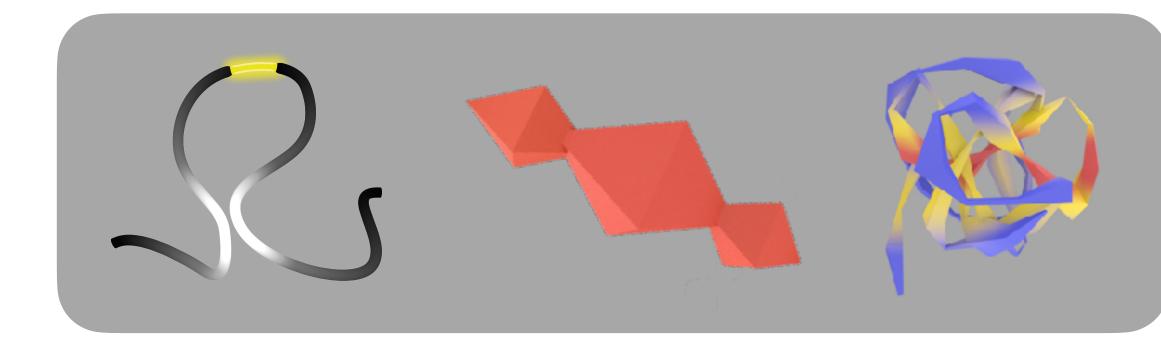
function (2D) RNA-seq · ChIP-seq · markers

distal

proximal



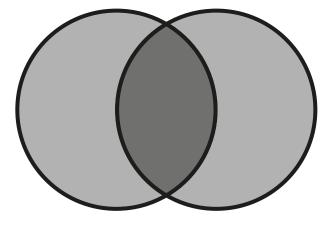
CTNNB1 ADIPDO	NEKRI TIR2 TIR3	્ય ન નનાય ના ન	E381 83 FOFE			, ∎
the definition of the second state of the second state of the second state of the second state of the second st effective states and the second state of the second st	and the sharehold and a state	الموالية برواد بتغلالهم ومعطاهم وتعاور أريبه	أوبر فأقفل الأمقية أردر والارتزاق وأخالهما	فيعاريه ومرافقاته والتقافين	فقدينين والتماد	والمقال مطال
in a second s	in the state of the	which the local second district of the	والمتحديقة والمحمولية	and an index of the second		ر الدائمين
(1) a transmission of the second sec second second sec	فيهاد والمعاط ومربيها يعامد ويقا	des alls en constituents de la constitue de la	ار الانطاب أفطالفيسيم وحوص الط	Read Minister Mail: Discost Mail, 18 Konner	والمتحد والطاط ومست	an ditana 20 20 20 20
el arcaneza llanar e caracellaren ar dan biter. Martin en elleren			en , and care marchine entry a	متعينا فليعاميهمات المارينية التناع	a and a state of the	20 



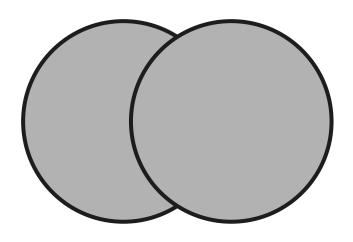


# Visual Grammar for 3DGenomics texture

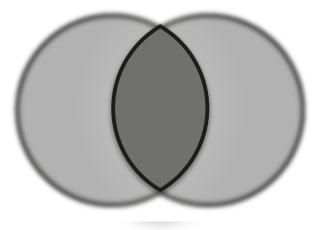




solid



glow + field depth





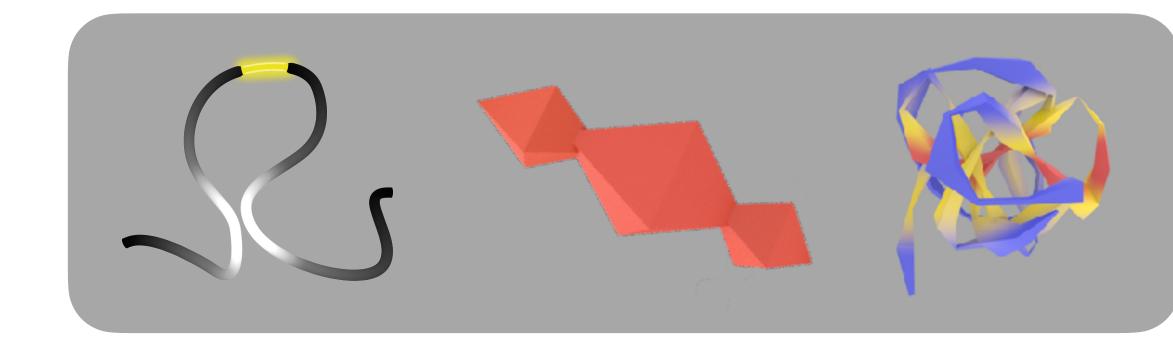
genes  $\cdot$  genome annotation



#### structure (3D) 3C-based data · 3D models



CTNNB1			b b b b b b	PIKAR1 II.4		ESR1	a FOFR ▶ ▶▶ ▶	N033	MYC  >  >  >  >	т: ( 44
44 4 KK			TERT		HLA-DEB1	4 4 4 4 4 4 4 4 6002		I BRAF PRKDG		4.4
and contraction participants	and the state of the state of the state			and a large start is still all a	بالصدر فتتعلقه وراقر	and the second states	وأربعة ورومه	and a state of the second second	and an and a second	
	hande en en serie de stête fait de ser ser References en serie de stête de serie de s									
										in the second second
a sin an a failtean a	Seattletine domained									إفرادهم
ald a ddi ann and llainnean a All a dda ann an Iallaith La	and the state of the							n an anna daoine Anna Anna Anna. Iomraithe anna an Anna Anna Anna		ور الاست. بار الاست.
	يريد والديد بالمراجع ومامر مري						ومرجلة بالسطا	ويرو والمادين وليمر مرهور		24
nii adda an in inii an a	المريدية <mark>،</mark> الماركية متراركي <mark>ة ا</mark> لأشيام. مريد الأمارة المريد مريد	And an a calledon of the calledon								ر الاندانينية. 24
						at the state of the		a sa ing aka di di kata sa sa	1. A Strategic States and the second	A back of the second



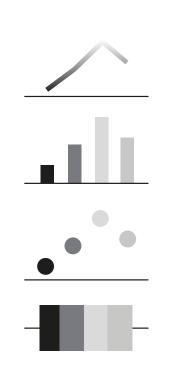


# Visual Grammar for 3DGenomics form

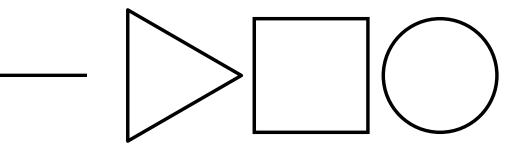


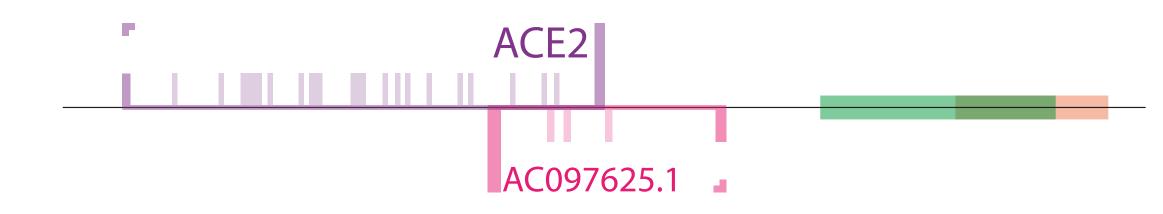
genes  $\cdot$  genome annotation



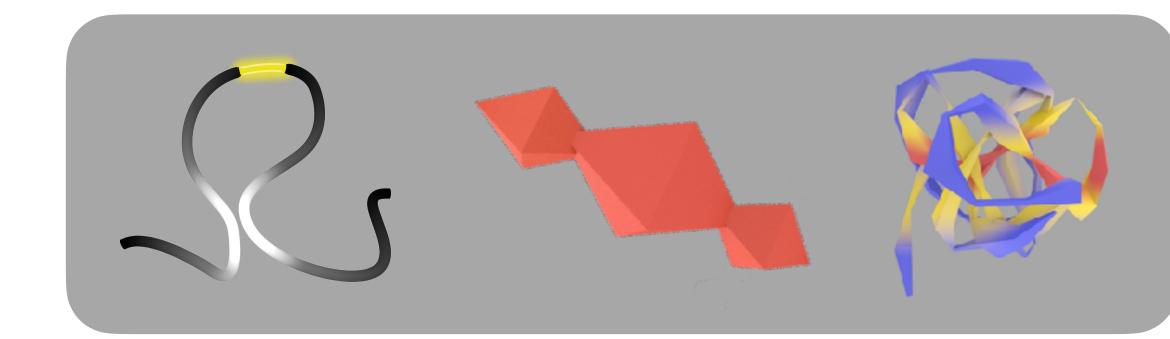








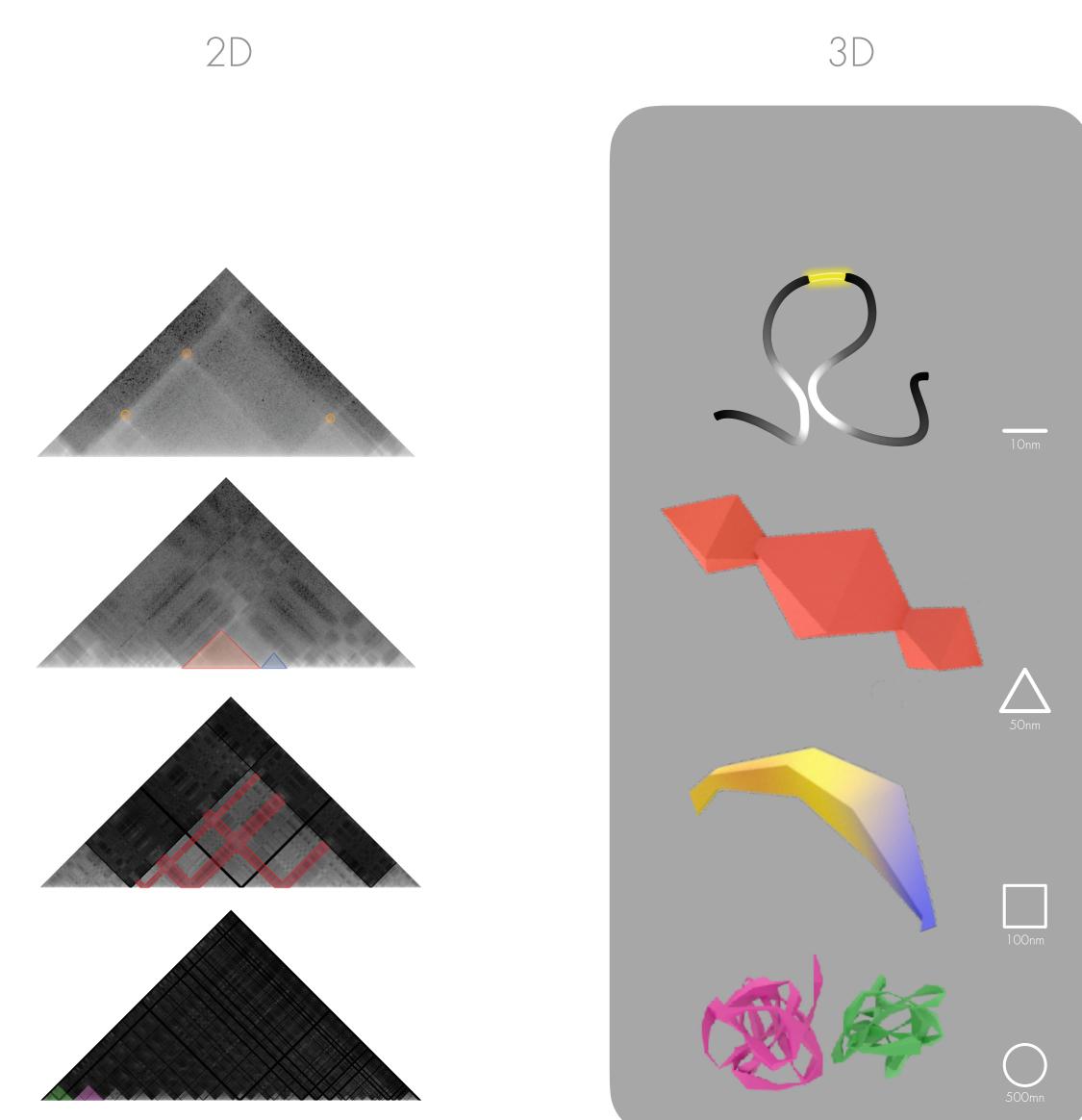
CTNNB1 ADIPOO ADIPO ADIPO ADIPOO A	NEKRI TIR2 TIR3			ND33		
n a baran da sa sa sa sa baran ya sa sa sa sa sa babina da sa sa A sa	in a standarditettin anadist Ana standarditettin anadist	ala an	enderlig de propie (n. 1916), en la de la berlik, en la ber Referie la de la constante de la bernet de la berliker en la s	a de campo de la composition de la comp	اللية التي يريد بين مع القديد ، ومرأة الألباط التي بدانية أوالا إذ رابطة	n da li
(b) the second linear second secon	A a b construction and a state of	an an ar an	ى ( ( ( ( ( ( ( ( ( ( ( ( ( ( ( ( ( ( (	andra da marcha da da da da da marcha. Andre a como de constan da marcha da	anda and dia ang ang ang ang ang ang ang ang ang an	المراسب ويسالد من ول الاسب
ومعاقبه والمعاوية والمتعدية والمستلك والمعالية	And the second beaution to		والمراجعة	من من الاستراكان عن المراجع من ماريد. مناطق المراجع ا		20 20 20 20

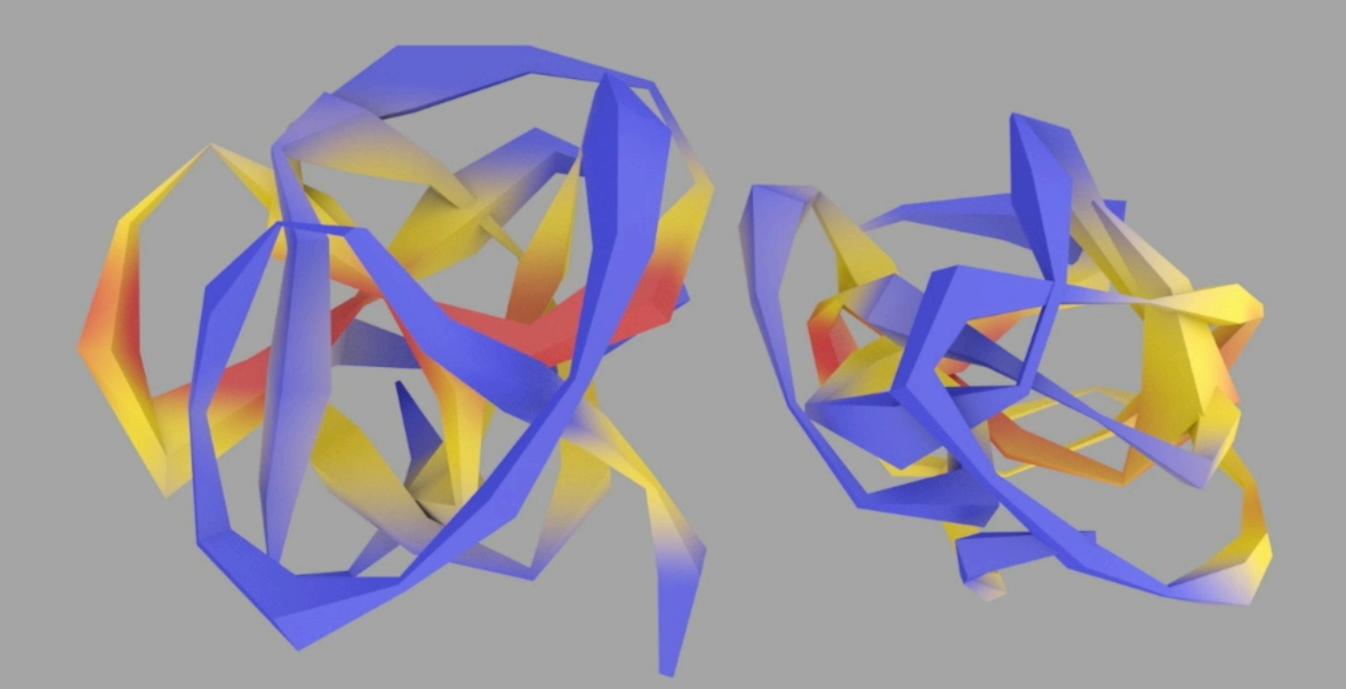




1D genes kb res ACE2 OOPS 10s kb res scale TADs 100s kb res \_ \_ \_ \_ \_ compartments Mb res chromosomes 1,0s Mb res

#### representation







## within a browser?

Gehlenborg Lab (Harvard U.)

Aiden Lab (Rice U.)

Search     Mapping and Sequencing     Genes and Gene Predictions     Phenotype and Literature     X   OMMI Alleles   Gene Expression in   33 TGGA Cancer Tissues   (CENCODE V23)   COSMIC Regions     LOVD Variants   Development Delays   Gene Interactions   OMNI Cyto Loci   Cin Gen   SNPedia     COVID 19     mRNA and EST   Expression   Novering on top of them.     Values     Add custom tracks			Enter location		×	chrX:15,56	60,138 -
Mapping and Sequencing   Genes and Gene Predictions   Plenotype and Literature   CoSMIC Regions   COSMIC Regions   CoSMIC Regions   COVID-19   mRNA and EST Expression   Expression   COVID-19   mRNA and EST Expression   Expression   This box explains object functions when hovering on top of them.	Q Search	X	8 - +				
Genes and Gene Predictions   Phenotype and Literature   X   OMNI Alleles   Gene Expression in   33 TCGA Cancer Tissues   (DENCODE v23)   COSMIC Regions   LOVD Variants   Development Delays   Gene Interactions   OMNI Cyto Loci   Clin Gen   SNPedia   CVID-19   mRNA and EST   Expression   Reset   Add outom tracks	Mapping and Sequencing					TI R2 TI R3	
OMNI Alleles   Gene Expression   (GENCODE v23)   COSMIC Regions   LOVD Variants   Development Delays   Gene Interactions   OMNI Cyto Loci   Cin Gen   SNPedia   COVID-19   mRNA and EST   Expression   Qrest   Add custom tracks	Genes and Gene Predictions		CHI TEHCA CERCIS	TNESE10 KL	4444 4 R SNCA 12	TERT	( ( Lanu - dalea
OMNI Alleles   Gene Expression   OSMIC Regions   LOVD Variants   Development Delays   Gene Interactions   OMNI Cyto Loci   Clin Gen   SNPedia   COVID-19   mRNA and EST   Expression   Qreset   Add custom tracks	Phenotype and Literature	×	d al angler de la fille de la		UT (Calendary) UT (Calendary)	tille and de the side of Side and a side of the side of	
Gene Expression in 33 TGA Cancer Tissues (GENCODE v23) COSMIC Regions LOVD Variants Development Delays Gene Interactions OMNI Cyto Loci Clin Gen SNPedia COVID-19 mRNA and EST Expression Reset Add custom tracks in This box explains object functions when hovering on top of them.	OMNI Alleles		distances allowers consider	las an <mark>tal</mark> tan in	condition in the line in the	فالم المحمد	a ana an
LOVD Variants Development Delays Cene Interactions OMNI Cyto Loci Cin Gen SNPedia SNPedia COVID-19 mRNA and EST Expression Rest Add custom tracks This box explains object functions when hovering on top of them.	33 TCGA Cancer Tissues		مراجع من من مراجع الأمر من مقر التي مراجع من من من الأمر من مراجع ال	ر میں میں میں اور	a selection and an an and	han in star o	ut union di
LOUD Variants     Development Delays   Gene Interactions   OMNI Cyto Loci   Clin Gen   SNPedia   COVID-19   mRNA and EST   Expression   Covering on top of them.	COSMIC Regions						
Gene Interactions   OMNI Cyto Loci   Cin Gen   SNPedia   SNPedia   COVID-19   mRNA and EST   Expression   Reset   Add custom tracks   This box explains object functions when hovering on top of them.	LOVD Variants						
OMNI Cyto Loci Cin Gen NPedia NPedia OVID-19 mRNA and EST Expression Reset Add custom tracks This box explains object functions when hovering on top of them.	Development Delays						
Clin Gen SNPedia SNPedia COVID-19 mRNA and EST Expression Reset Add custom tracks This box explains object functions when hovering on top of them.	Gene Interactions						
SNPedia COVID-19 mRNA and EST Expression  Reset Add custom tracks  This box explains object functions when hovering on top of them.	OMNI Cyto Loci						
COVID-19 mRNA and EST Expression Reset Add custom tracks This box explains object functions when hovering on top of them.	Clin Gen						
<ul> <li>mRNA and EST</li> <li>Expression</li> <li>Reset Add custom tracks</li> <li>This box explains object functions when hovering on top of them.</li> </ul>	SNPedia						
Expression Reset Add custom tracks This box explains object functions when hovering on top of them.	COVID-19						
<ul> <li>Reset Add custom tracks</li> <li>This box explains object functions when hovering on top of them.</li> </ul>	mRNA and EST						
This box explains object functions when hovering on top of them.	Expression						
hovering on top of them.	Reset Add custom tracks						
	(i) This box explains object functions when hovering on top of them.						

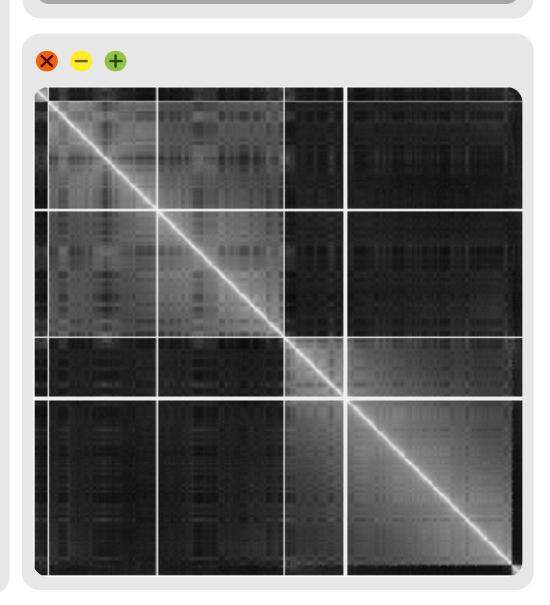
#### http://higlass.io · http://aidenlab.org/juicebox/ · http://aidenlab.org/spacewalk/ · http://3DGenome.org · http://vis.nucleome.org

× - +

Yue Lab (Northwestern U.)

Ma Lab (Carnegie Mellon U.)

—15,602,945			
14 TNF • • • • • • • • • • • • • • • • • • •	Fari la forr Nos		
નનાત ના ના ના ના	ય સંચય ને તેલના તેના	નાનાનાં નનાન નાનના	4 44
NRSC HLA-DRB1	SOD2 ABCB1 BRAF	PRKDO ODKN2	10+6
an e dining the state of the second state of the	لدور ويعقدون والدر أأقط وطوقان وا	يدعدنها للحدر مطافعة والمعويد فت	
a dha an	n an	and a strength of the first of the strength of	
and a state of the		a the tax is a set of the set	
the second se	والاعتباد والمطاوية ومطالب الطاعية ومحاجب بالتها فتعقص	ورجاء بتنفا باللهش خانقتين وراسا وروا الارات التفاري	الأفعالي المريدية
		a de allais de processo, comba contribuira de com a característica de processo como transferidade en el t	and and and
1 1 4 1	carrantel construction descriptions de acceptional construction descriptions des		1
4.1		a scalar la nacional contra contra la secon contra ta provincio con contra la secon contra la contra contra contra da secon	20+0 20+0
4.1	carrantel construction descriptions de acceptional construction descriptions des	and the second	20+6 20+6 20+6









Yolanda Justicia Contreras ELISAVA



Clementina Altube ELISAVA



Carla Molins Pitrach ELISAVA



Francesc Ribot Puntí ELISAVA

#### Elisava Research









Santiago Bonet Garrote ELISAVA



Jonathan Chacón Pérez ELISAVA



Mike Goodstadt CNAG-CRG

