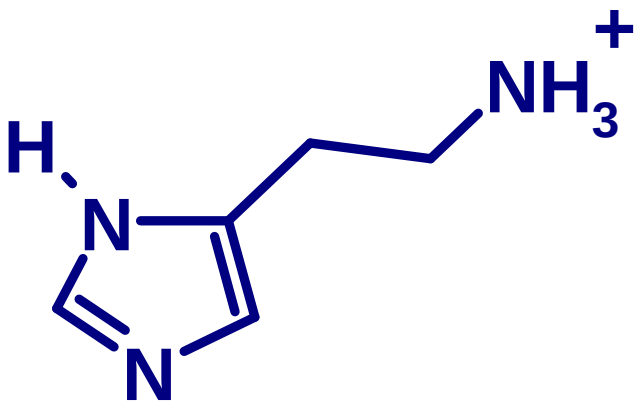
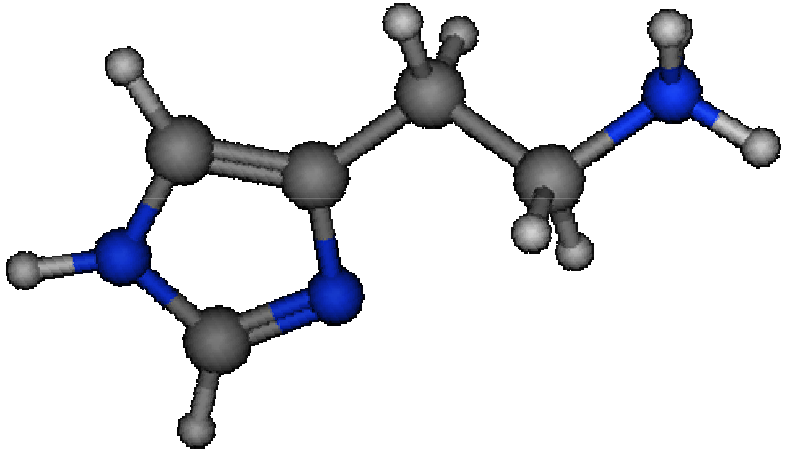
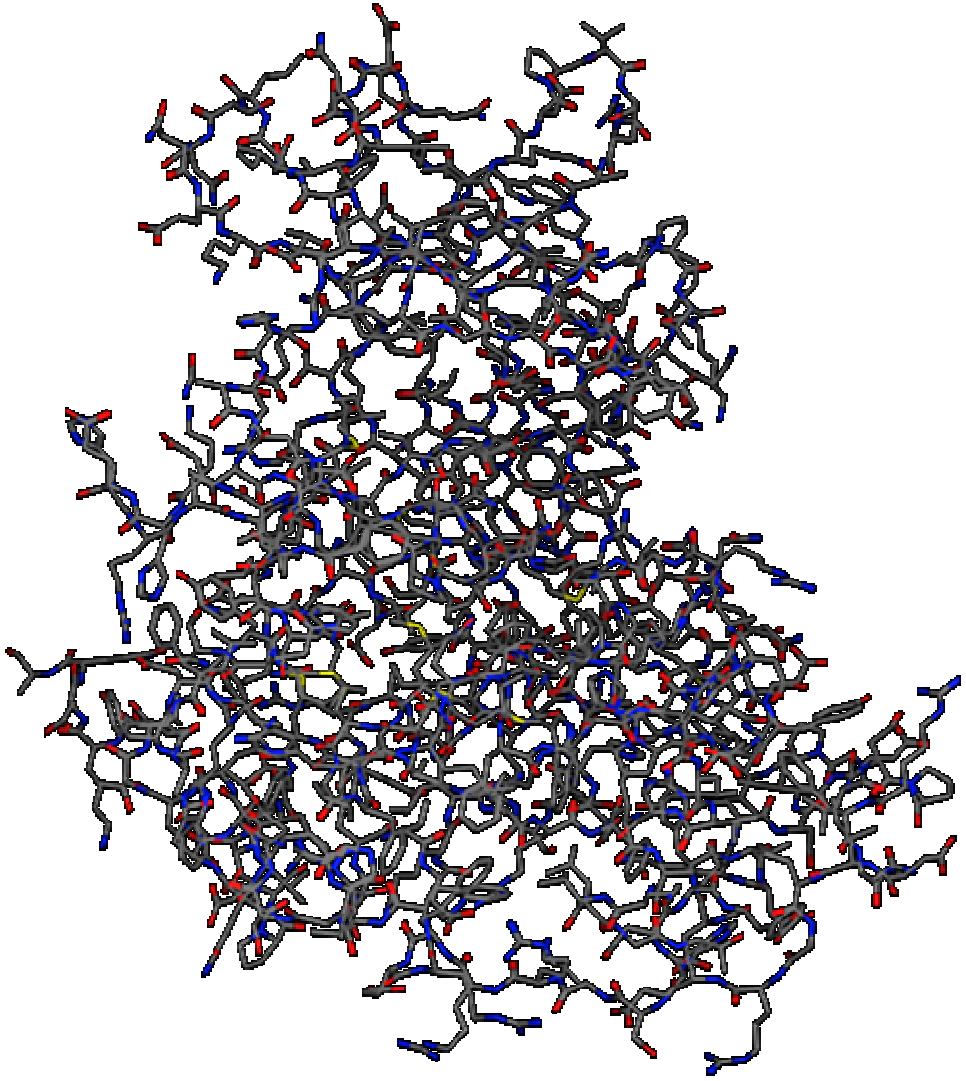
Now you understand how crucial is the 3D representation of a molecular structure:



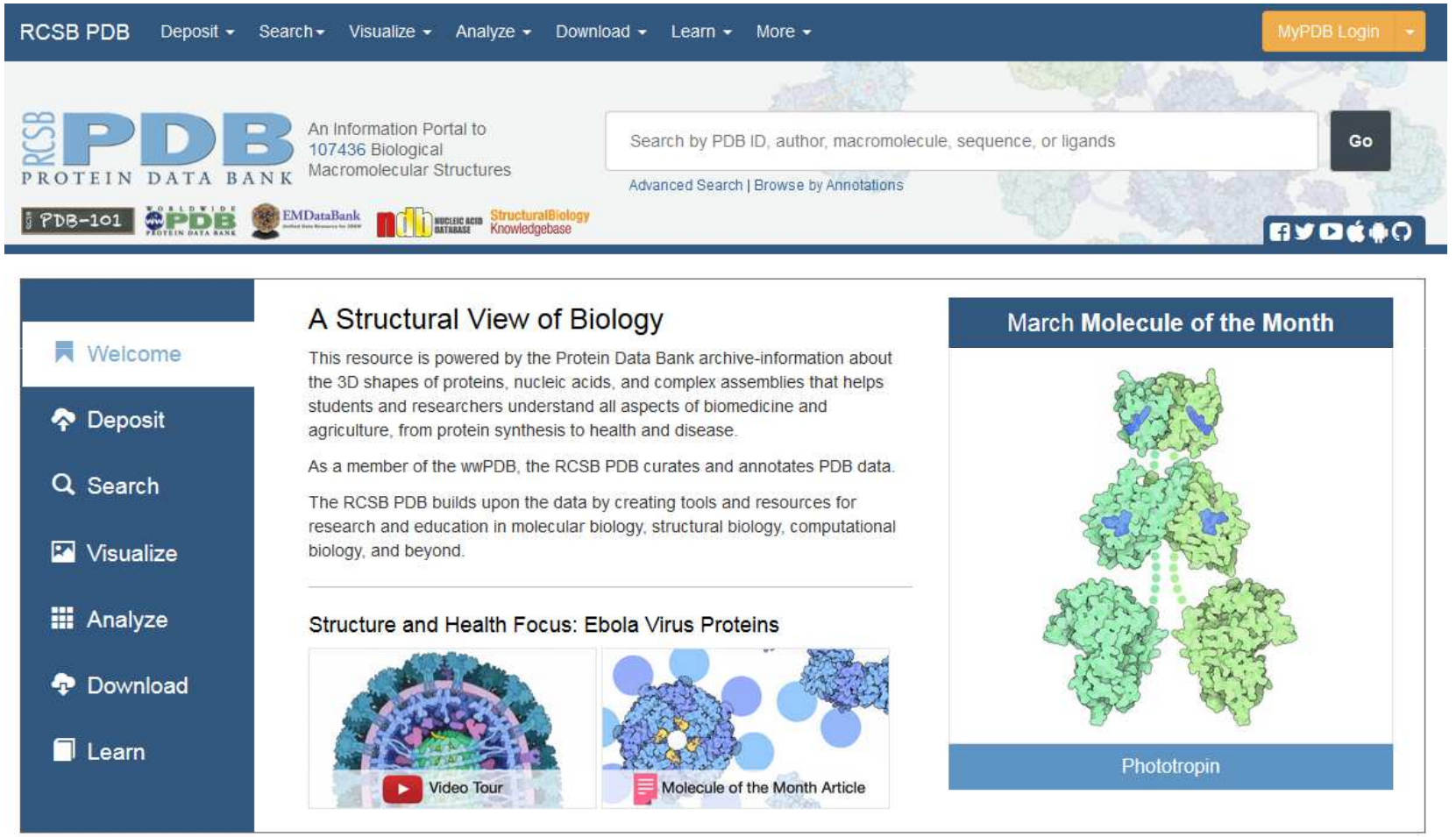


… and specially this one:



… and this is our favorite hunting place!

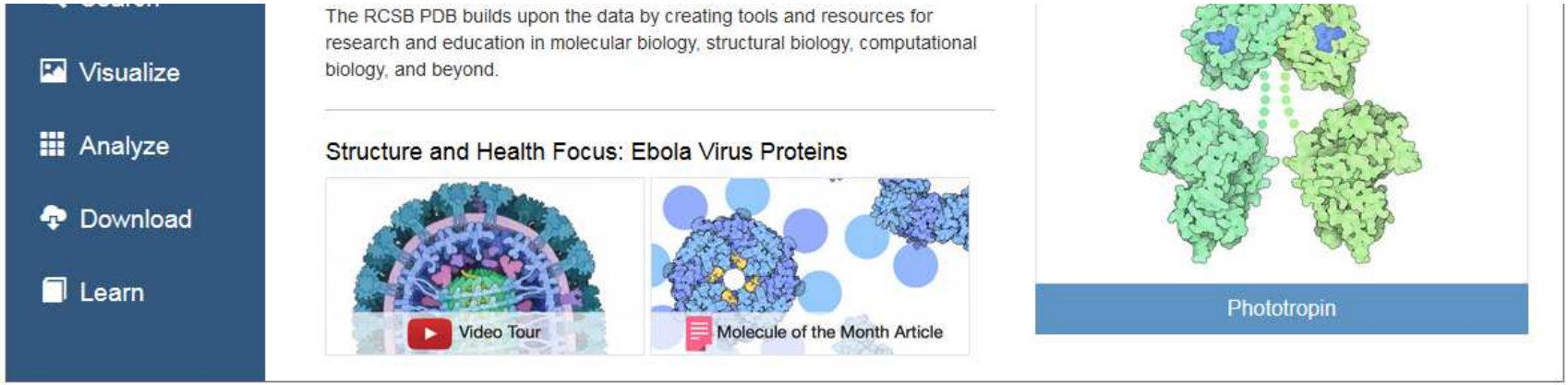
www.rcsb.org



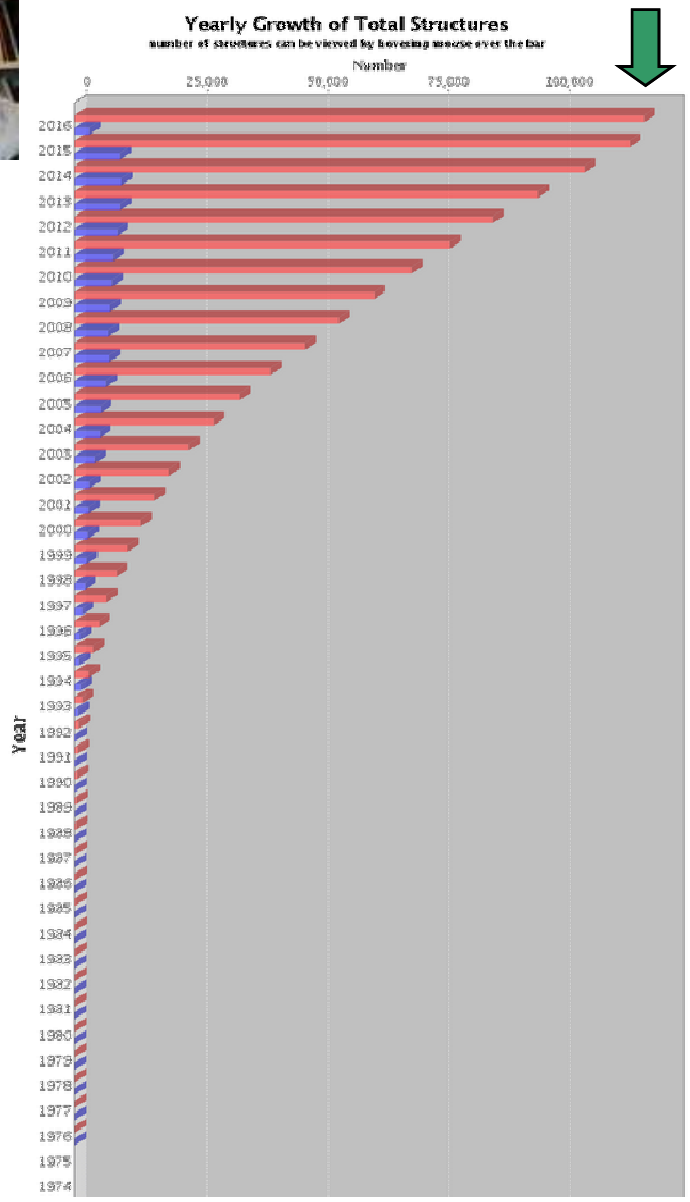
PDB… in numbers:

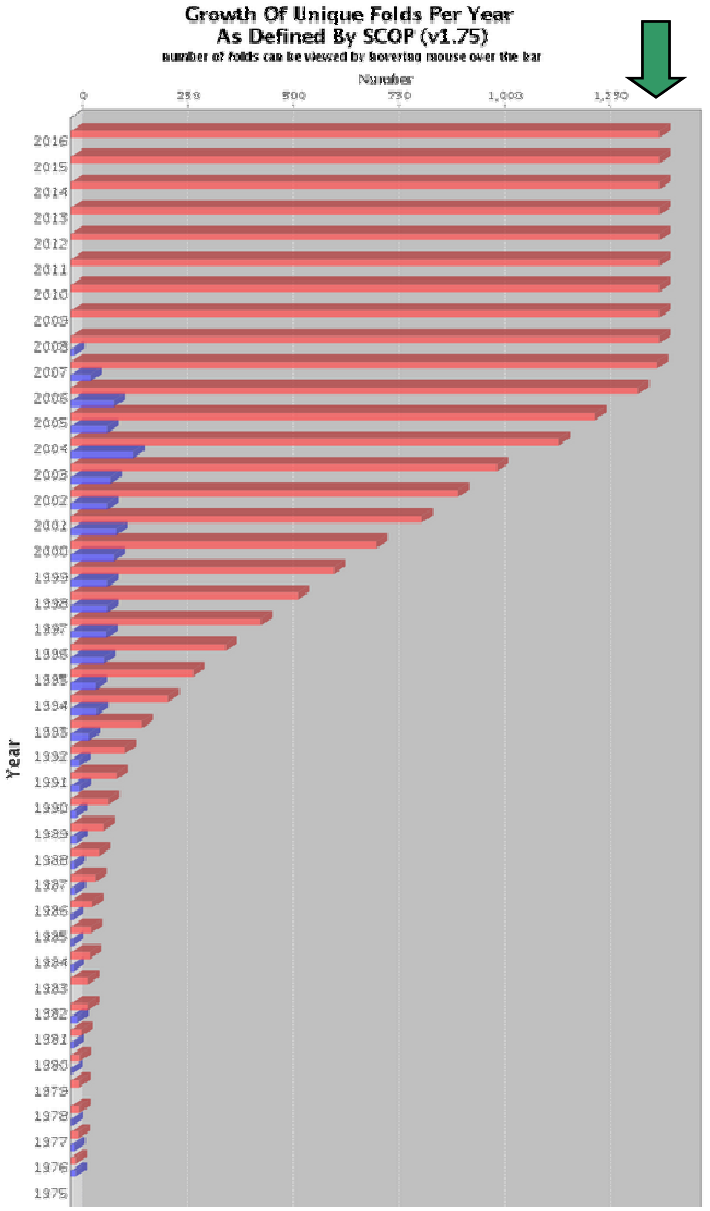
RCSB PDB Deposit Search Visualize Analyze Download Learn More MyPDB Login

Exp.Method	Proteins	Nucleic Acids	Protein/NA Complexes	Other	Total
X-RAY	98341	1713	4979	4	105037
NMR	9953	1140	231	8	11332
ELECTRON MICROSCOPY	715	29	249	0	993
HYBRID	87	3	2	1	93
other	173	4	6	13	196
Total	109269	2889	5467	26	117651



PDB... in numbers:



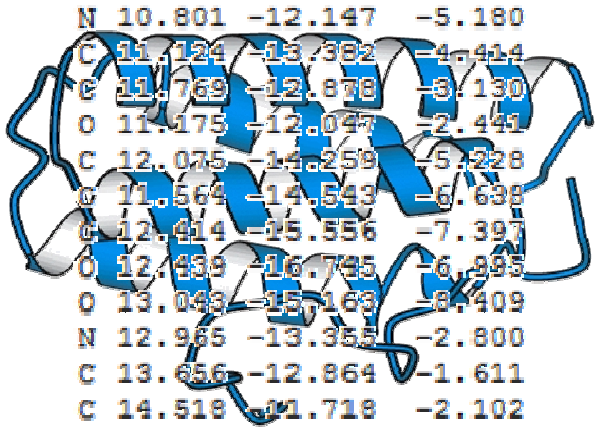


07 April 2016

My best friends…

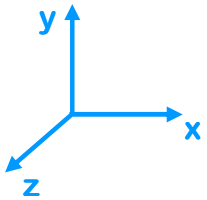
NMR
Spectroscopy

X-Ray
Crystallography

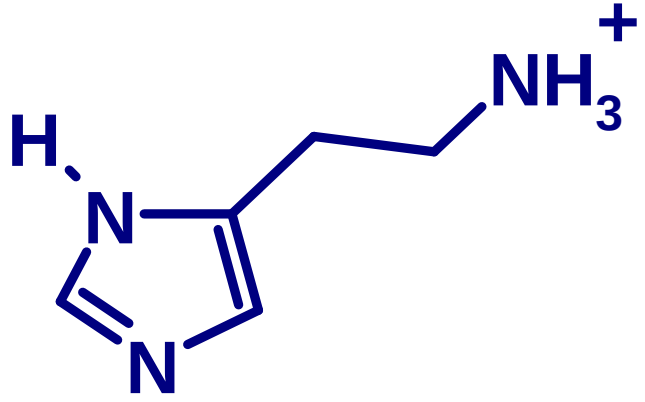


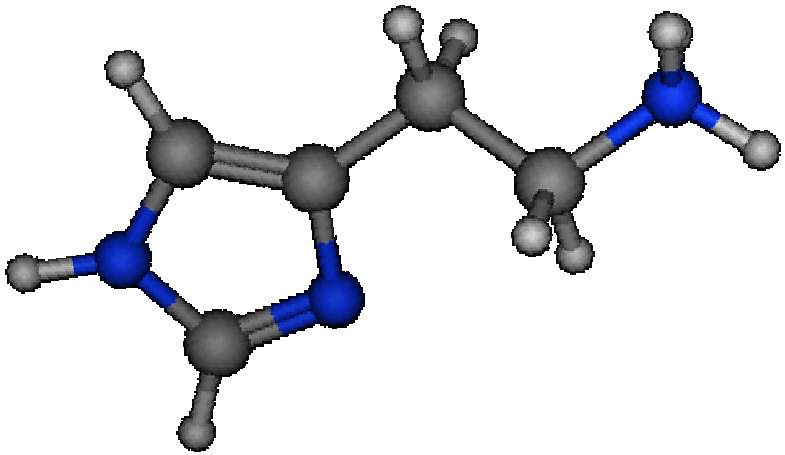
Comparative/Homology
Modeling

3D



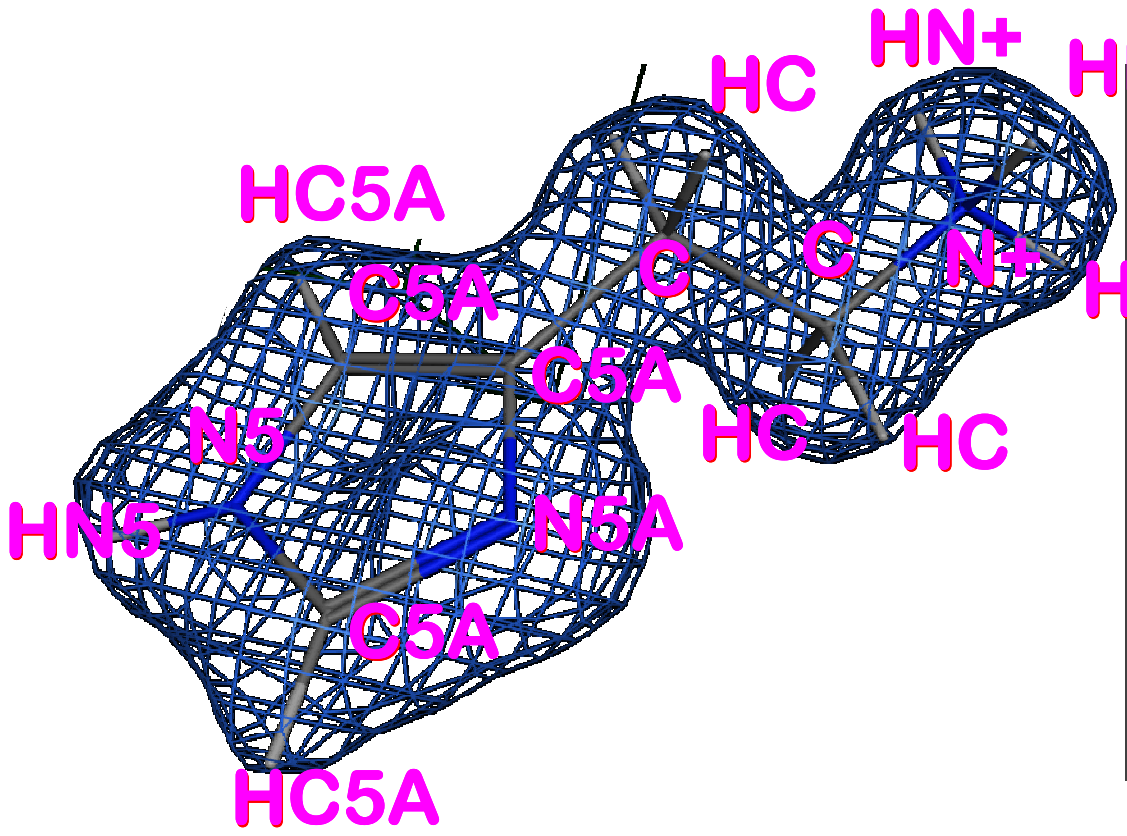
First of all… let's convert!





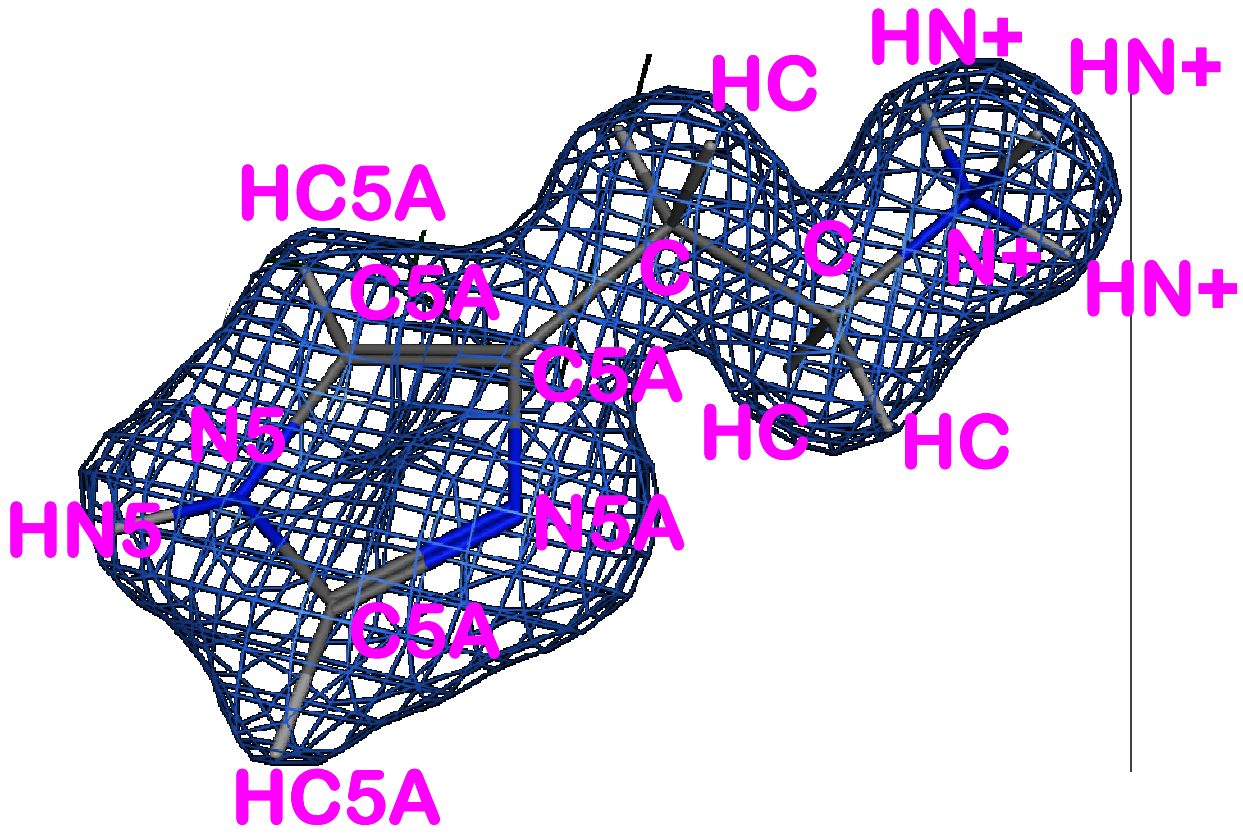
How? But learning from crystallographer, of course!

1. From what we observe, we learn.



HN+	ATOM	ATOM-TYPE
	H	HN+
	N	HC
HN+	C	HC5A
		N+
		N5A
		N5
		C
		C5A

1. From what we observe, we learn.



DISTANCES (Å):

HN+ - N+ = 1.03

HC – C = 1.10

HC5A – C5A = 1.08

HN5 – N5 = 1.01

N+ - C = 1.46

C – C = 1.52

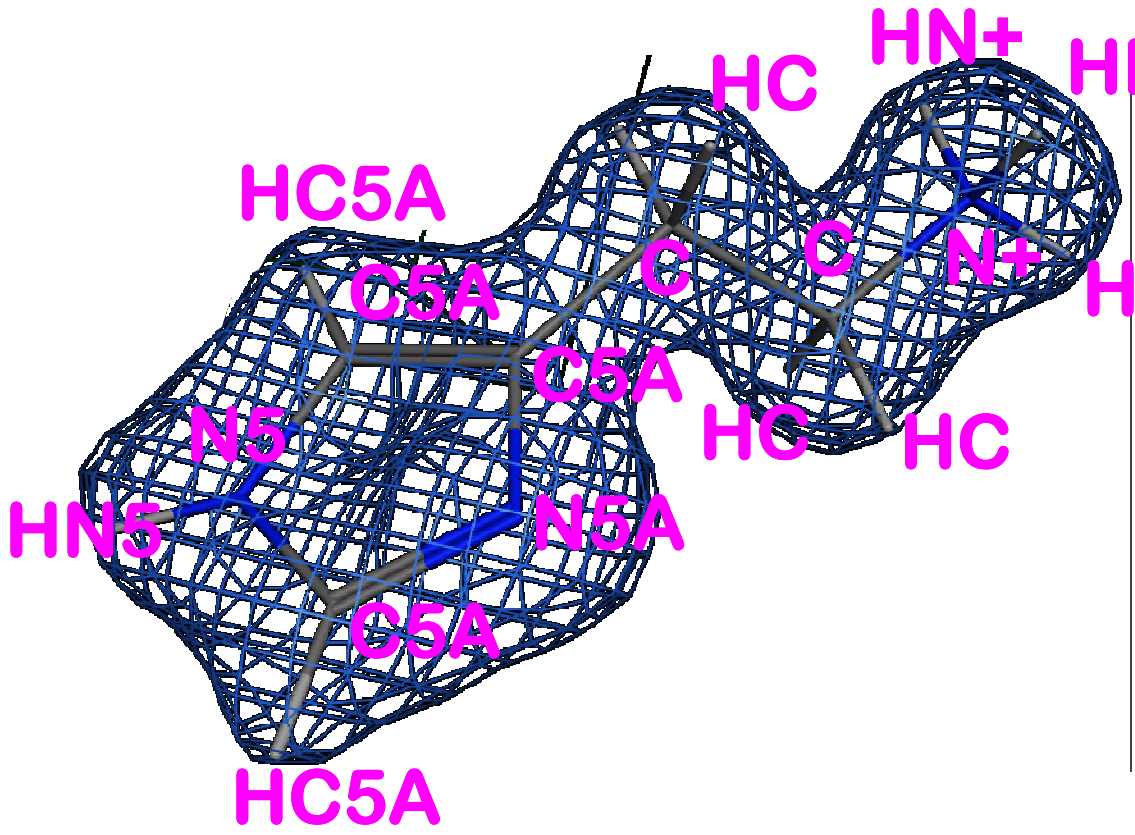
C – C5A = 1.51

C5A – C5A = 1.38

C5A – N5A = 1.35

C5A – N5 = 1.35

1. From what we observe, we learn.



ANGLES (°):

HN+ - N+ - HN+ = 107.1

HN+ - N+ - C = 111.8

HC - C - N+ = 108.2

HC - C - C = 111.4

C - C - C5A = 111.5

C - C5A - C5A = 126.2

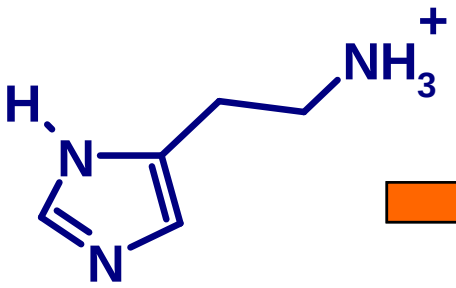
C5A - C5A - HC5A = 130.5

C5A - C5A - N5 = 107.2

C5A - N5 - C5A = 108.5

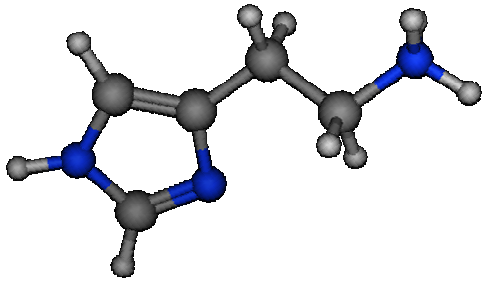
N5 - C5A - N5A = 108.7

2. … and ones learned, we can repeat!



ATOM- TYPES

GEOMETRIC PROPERTIES TABLE



it will be a good conformation?

a simple MMS tool…

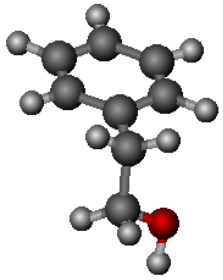
http://mms.dsfarm.unipd.it/VirtualLab/LigandBuilder.php

Wellcome to the Ligand Builder

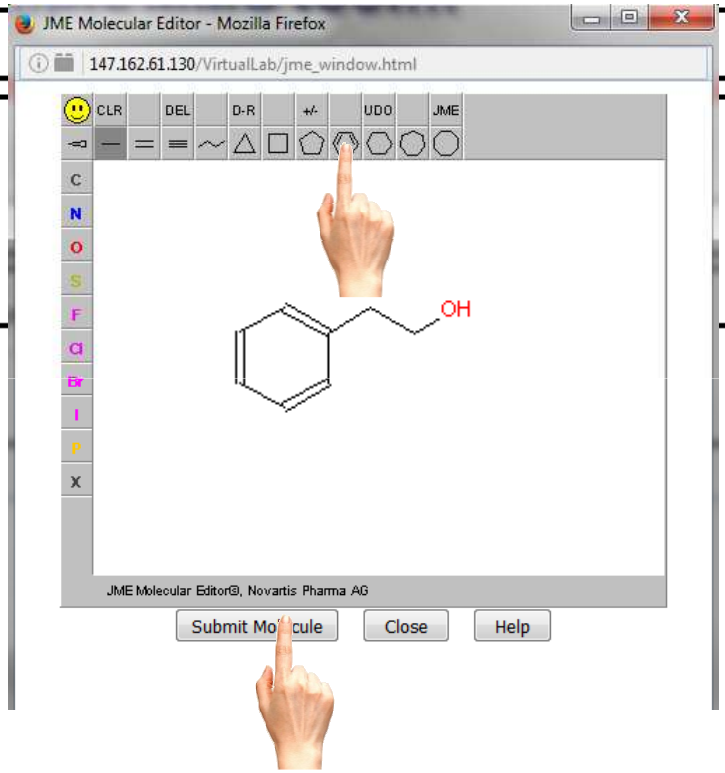
Please enter here a valid SMILES

OCCc1ccccc1 | Clear | Draw Molecule

UN | PDB Download | SDF Download



Molecular Weight	122.166997
Mass density	0.6559283482
Octanol/water log P	1.22137
Water solubility log S	-1.21737
Surface Area	143.6158526
Polar Surface Area	20.23
vdW volume	186.2505219
H-bond acceptors	1
H-bond donors	1
Hydrophobes	8
Acidic atoms	0
Basic atoms	0
Aromatic atoms	6
Rotatable bonds	2
Chiral centers	0
Druglike	1
Leadlike	1
Reactive	0



Exploring conformational space…

What we are still orphans:

- Virtualize molecular topology (shape and volume);

- **Virtualize the generation of alternative conformers;**

- Virtualize the evaluation of the stability of each conformer.

Why we need it?

- Identification of most probable 3D arrangements of a molecule in a gas or solution phase;

- Several chemical properties is conformation-dependent;

- Any kind of complementarity (topological, interactive…) is conformation dependent.

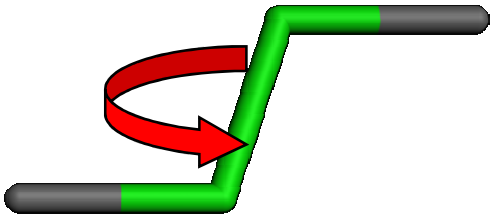
Do you remember these two concepts:

Stability **as a measure of the geometrical deformability of an object;**

Rigidity **as a measure of the reduction degree of the geometrical deformability of an object.**

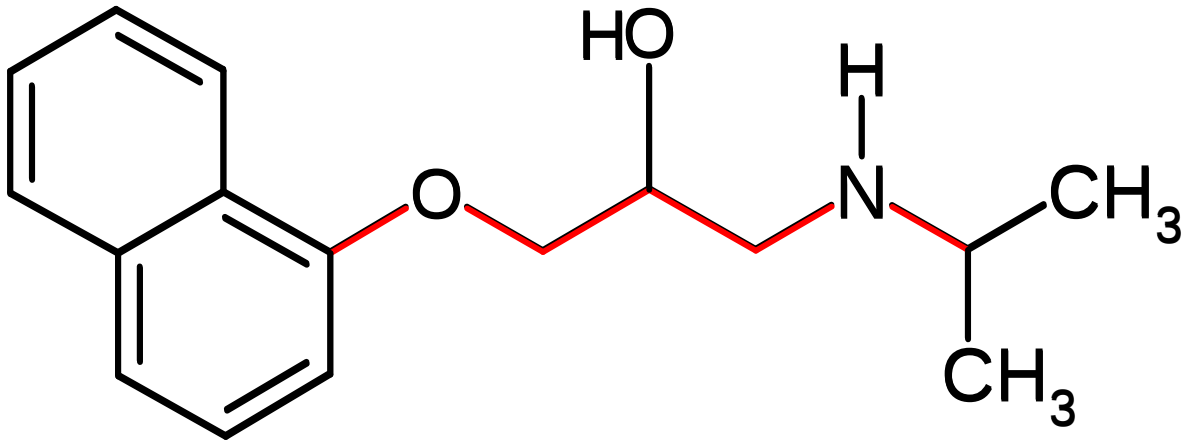
An easy way to determine *molecular rigidity*.

A ***rotatable bond*** is defined as any single non-ring bond, attached to a non-terminal, non-hydrogen atom. ***Amide C-N bonds are not counted because of their high barrier to rotation.***



… and it is easily countable!!!

Just count it:



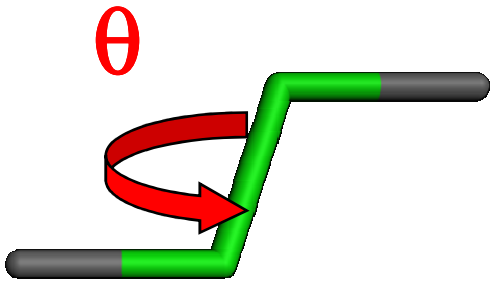
6

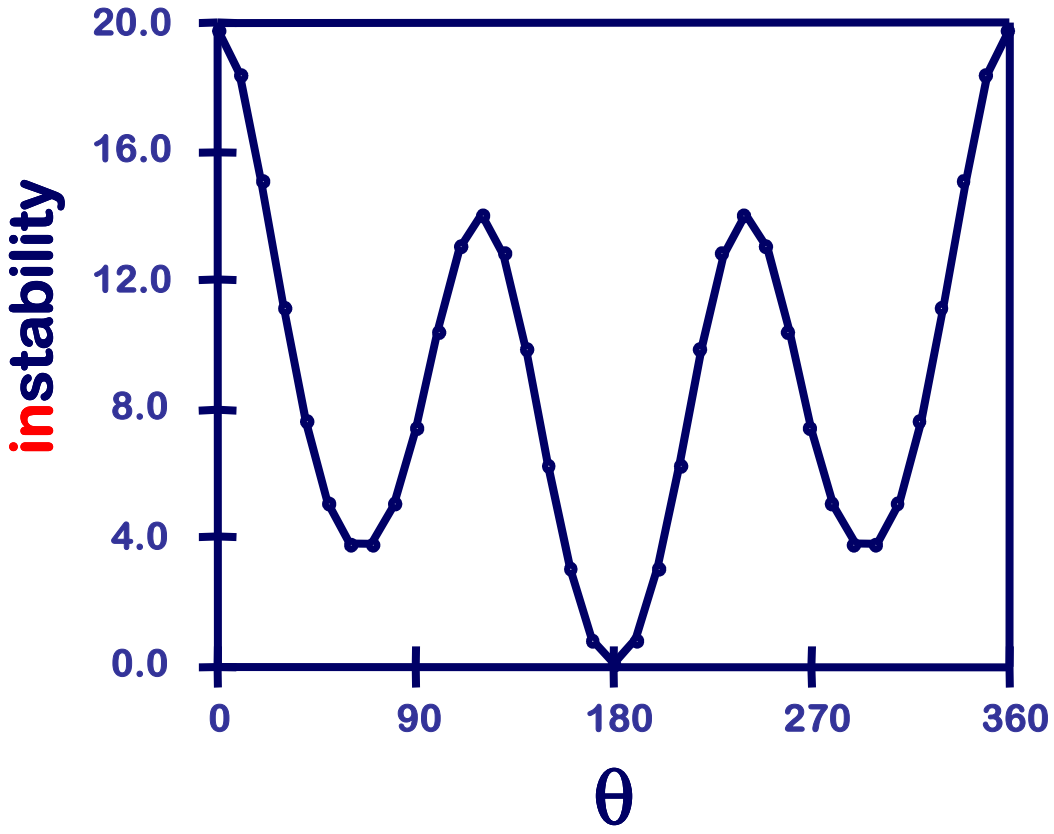
Do you remember these two concepts:

Stability **as a measure of the geometrical deformability of an object;**

Rigidity **as a measure of the reduction degree of the geometrical deformability of an object.**

Few fundamental concepts:

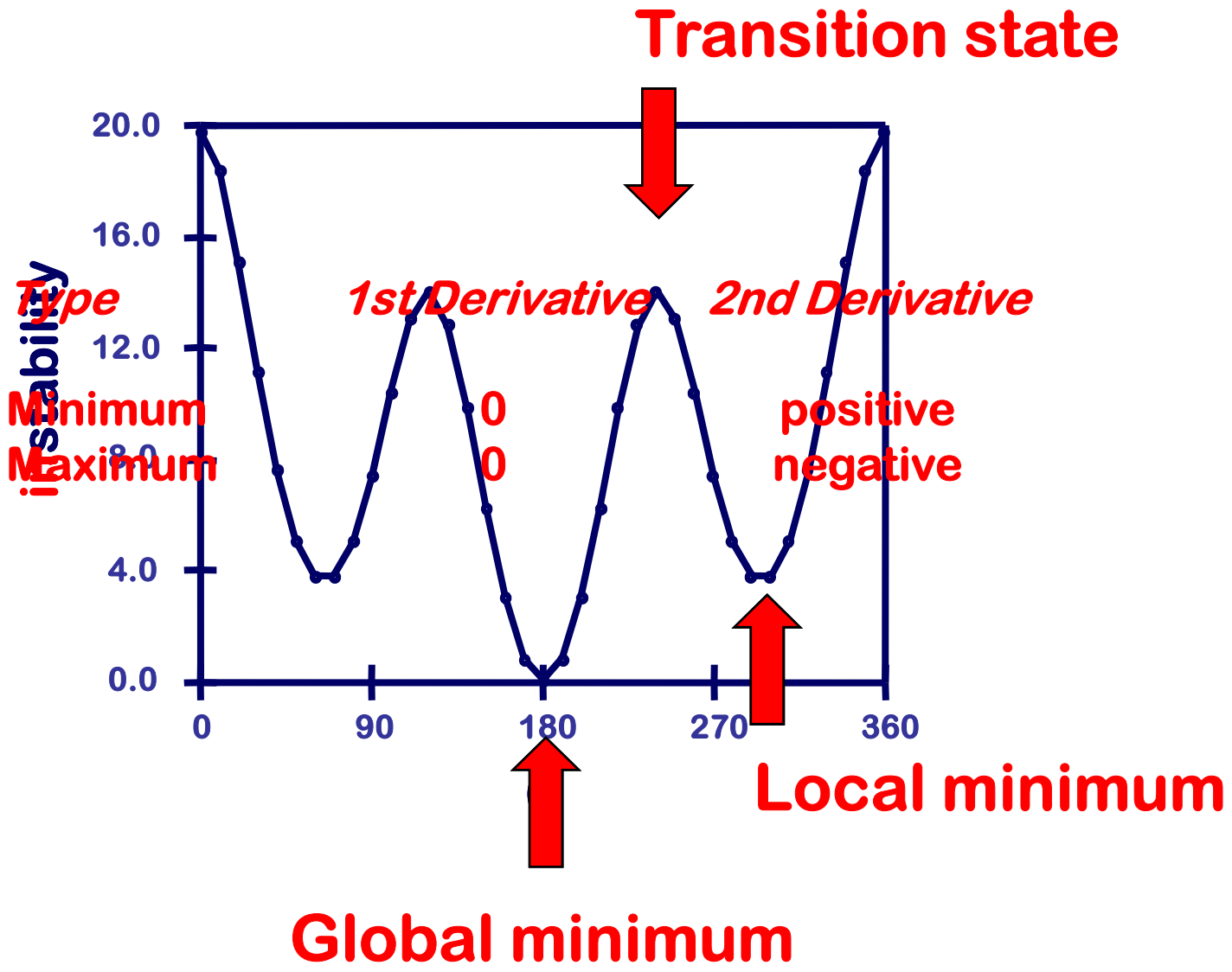




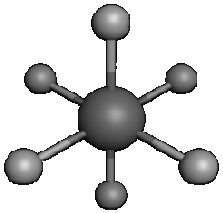
Conformers:

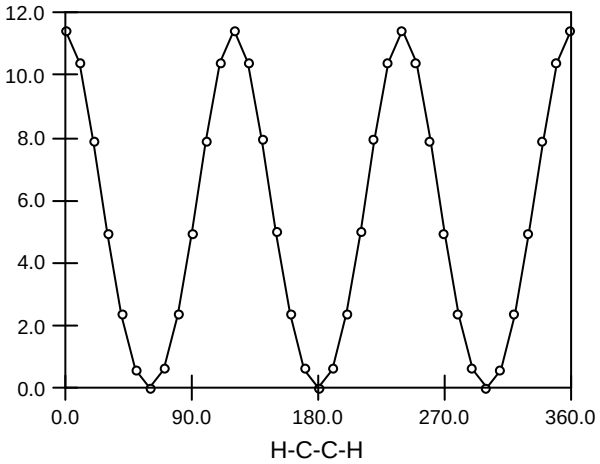
Structures differing only by rotation around one or more rotatable bonds.

Few fundamental concepts:

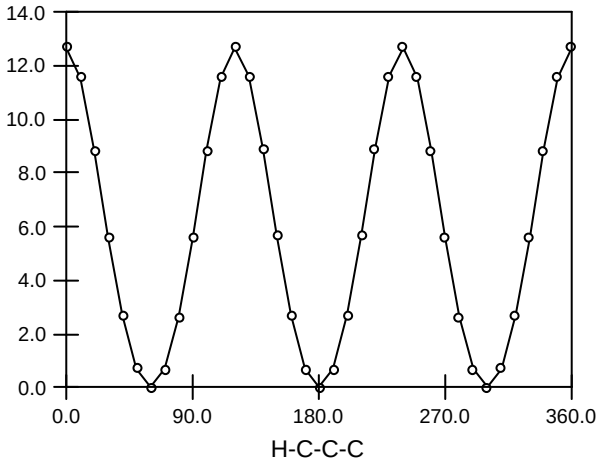


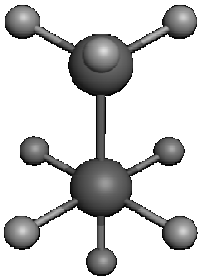
For a molecular system with N rotatable bonds, there are 3^N potential minima.



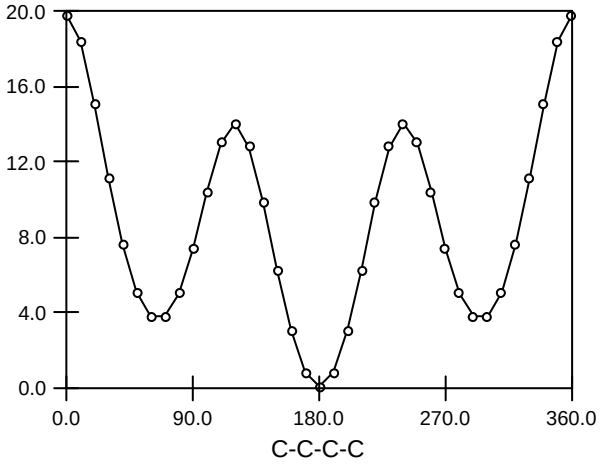


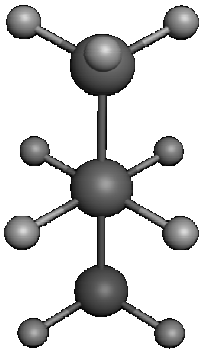
ethane





propane

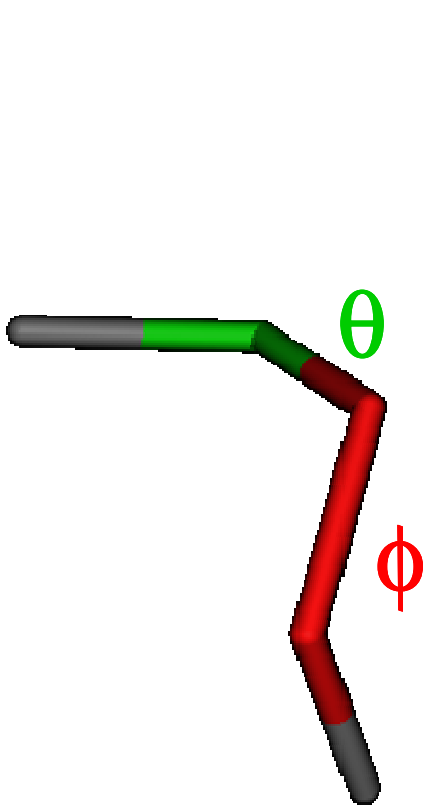


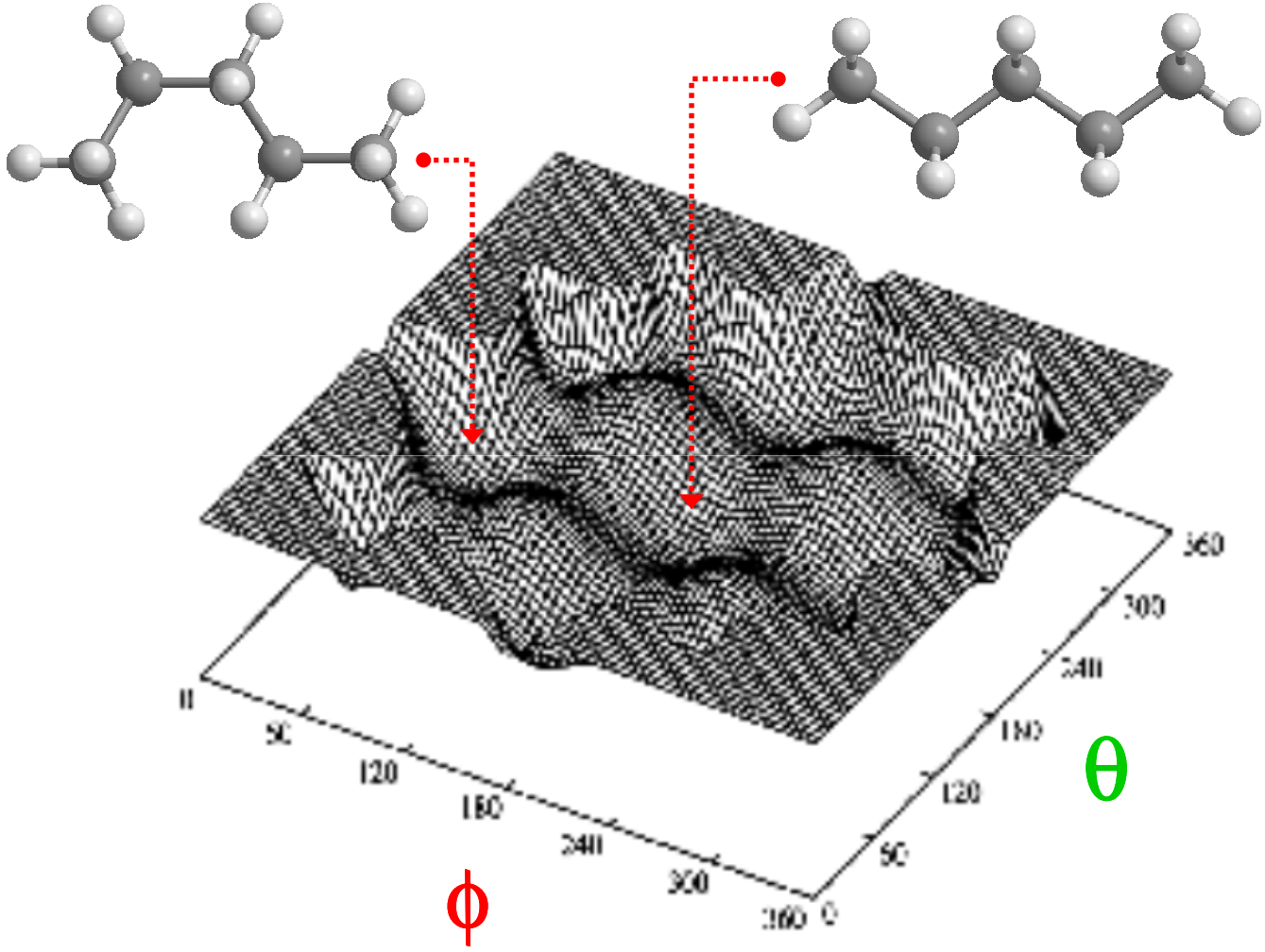


butane

N = 1

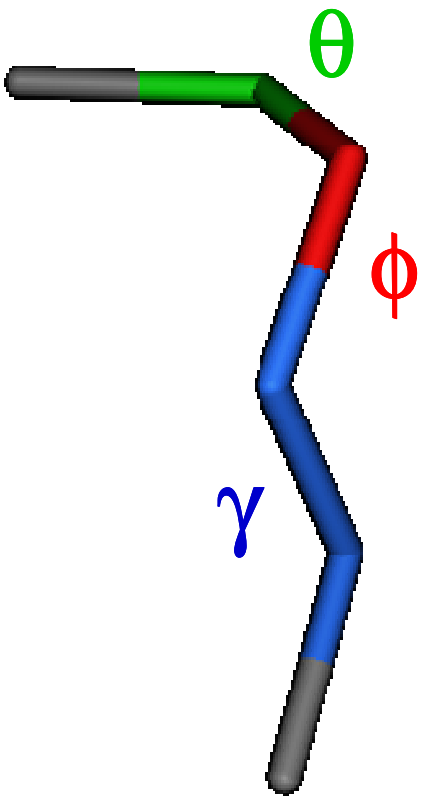
For a molecular system with N rotatable bonds, there are 3^N potential minima.





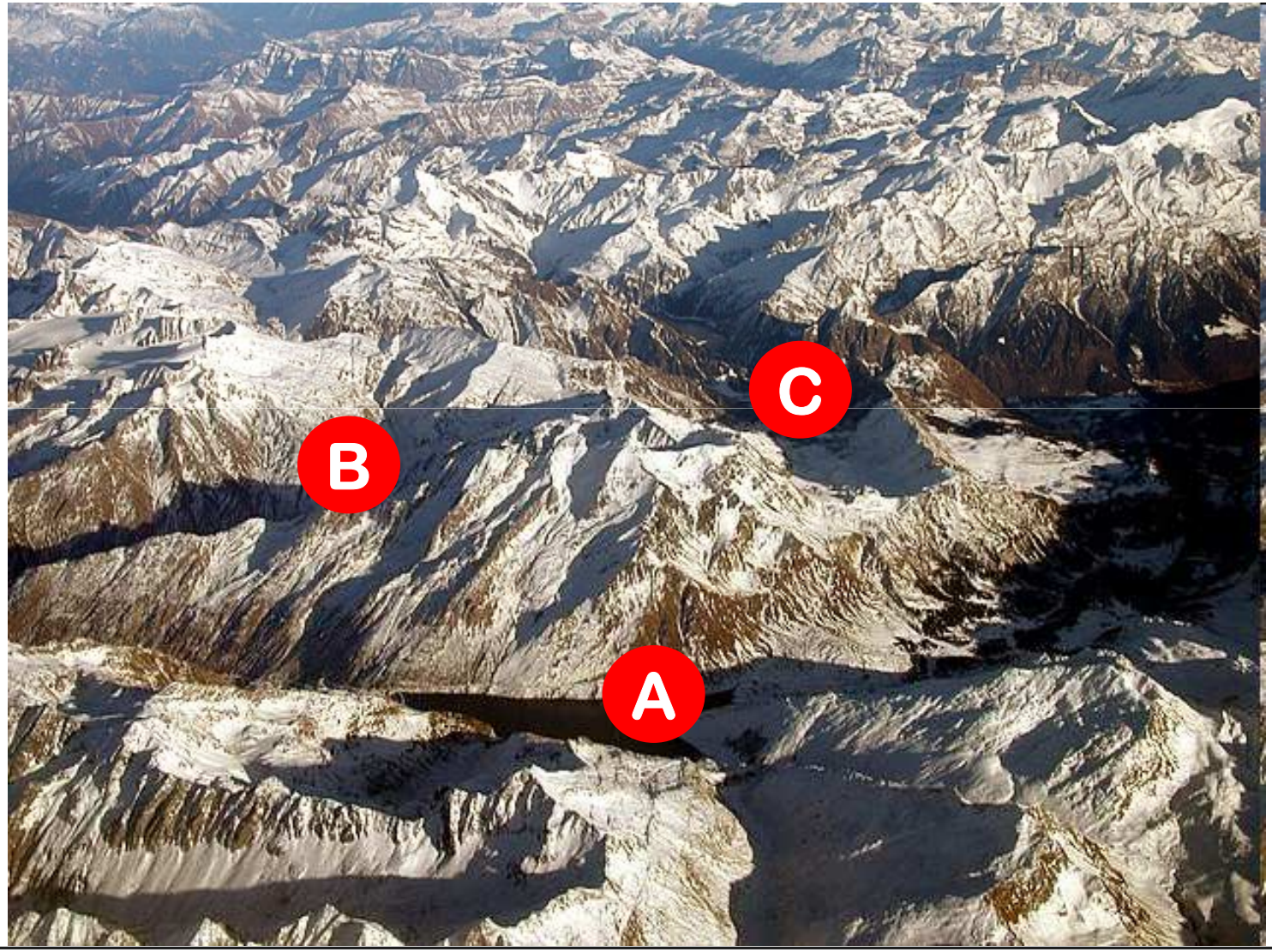
N = 2

For a molecular system with N rotatable bonds, there are 3^N potential minima.



N = 3

How we can explore conformational spaces:

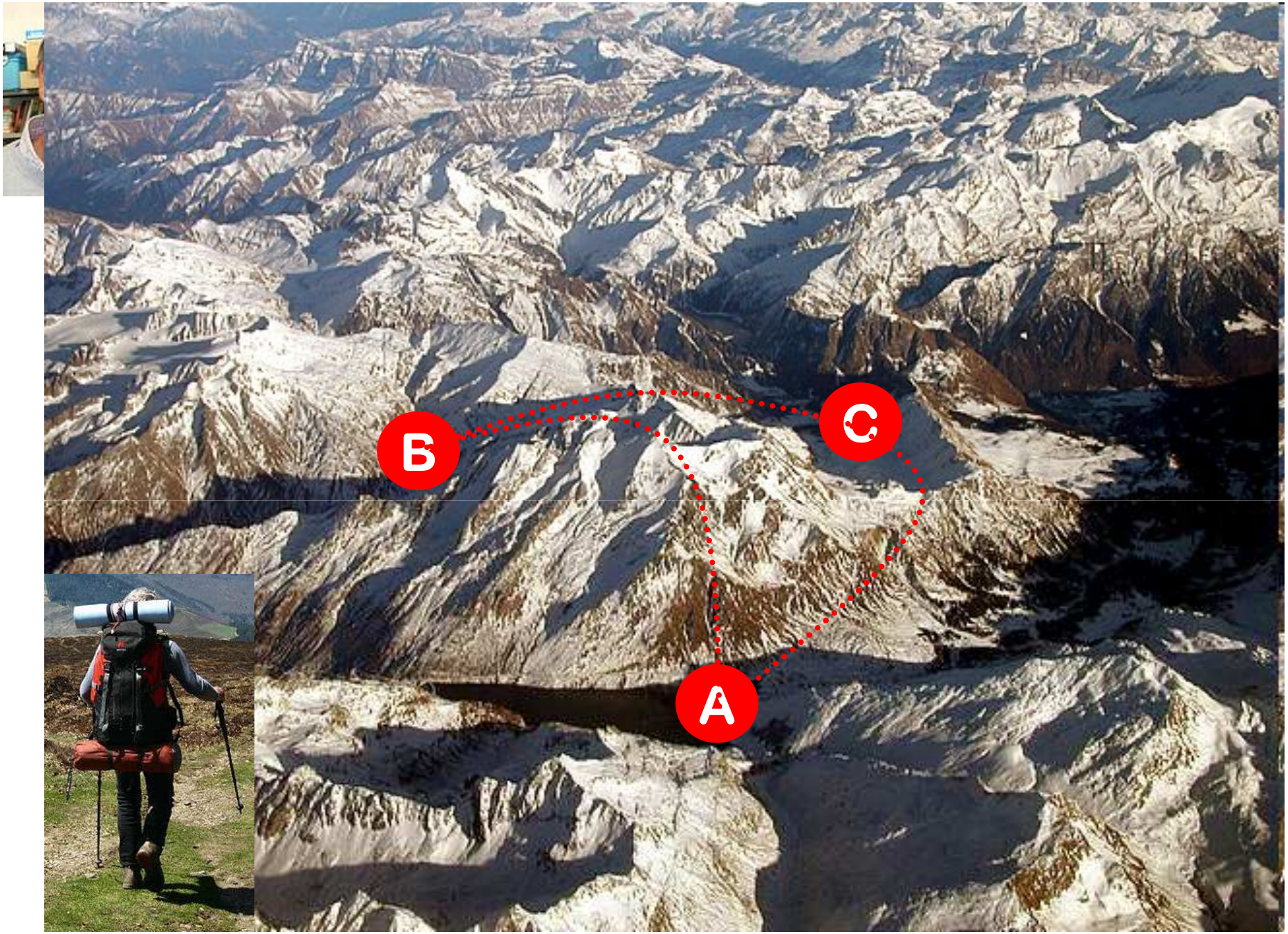


How we can explore conformational spaces:

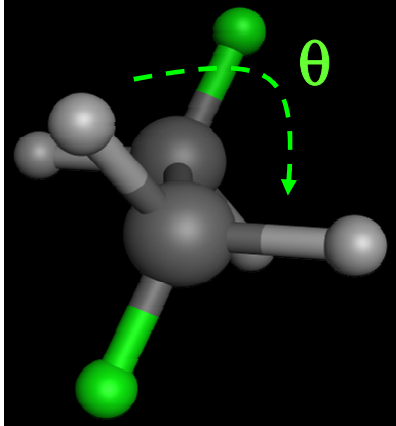
1. Following time evolution of a chemical structure (molecular dynamics);

2. Geometrical mapping of the conformational space.

Systematic analysis;
Stochastic methods (Random Search);
Genetic algorithms;
Distance Geometry method.


B
C
A

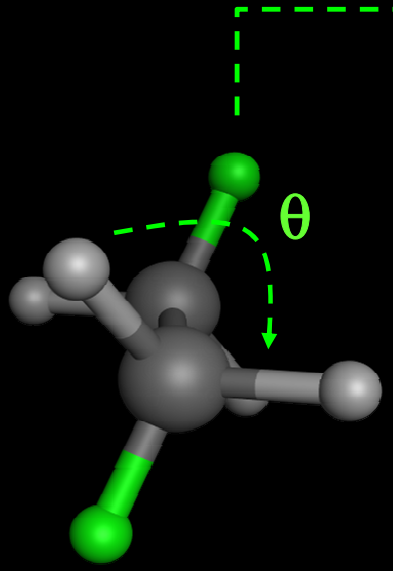
Algorithmically is very clear…

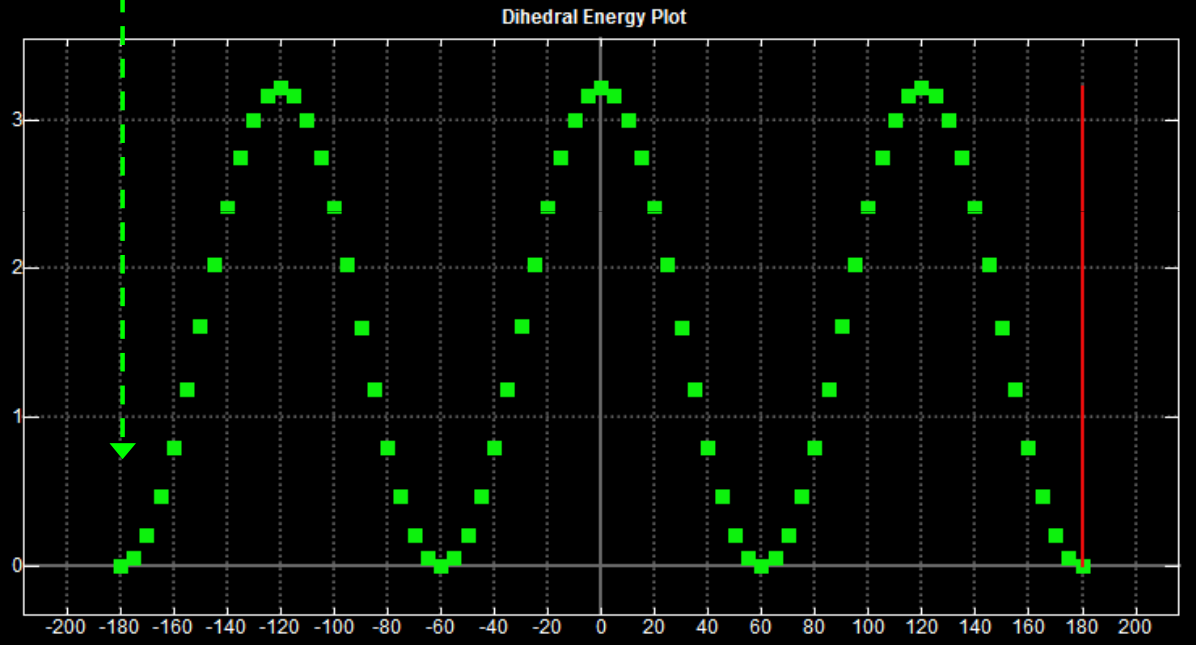


For a randomly chosen dihedral angle, systematically rotate about that bond in discrete angular increments.

If we have a measure of structure stability:

stability equation





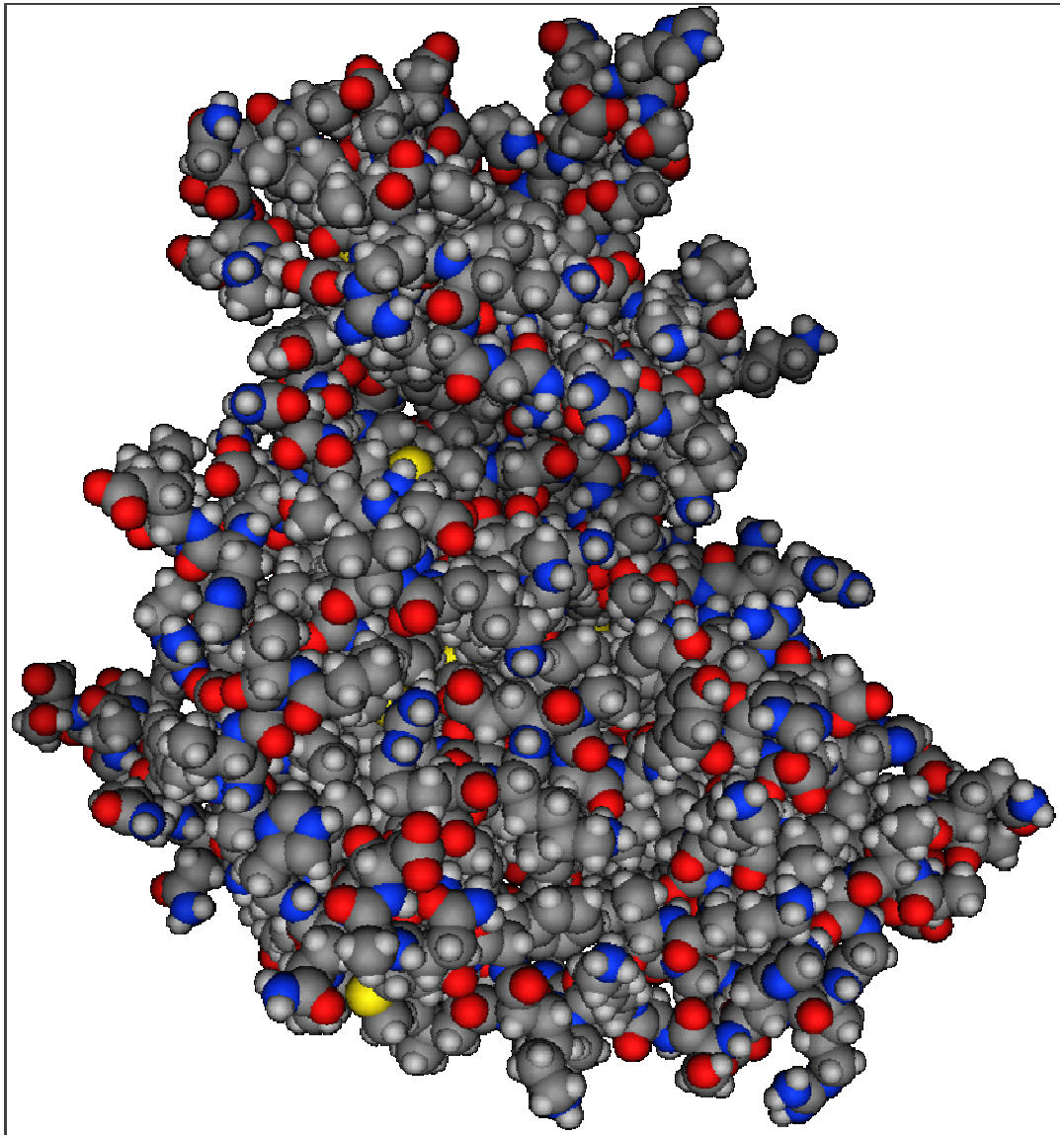
But we have a very severe limitation…

For a combinations of dihedral angles:

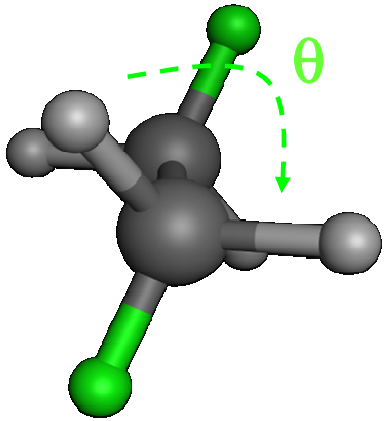
Rotatable bonds	θ	n_{conf}	time
3	30	1728	…
5	30	248832	…
7	30	36×10^6	…

This phenomena is named COMBINATORIAL EXPLOSION.

But we have a very severe limitation…

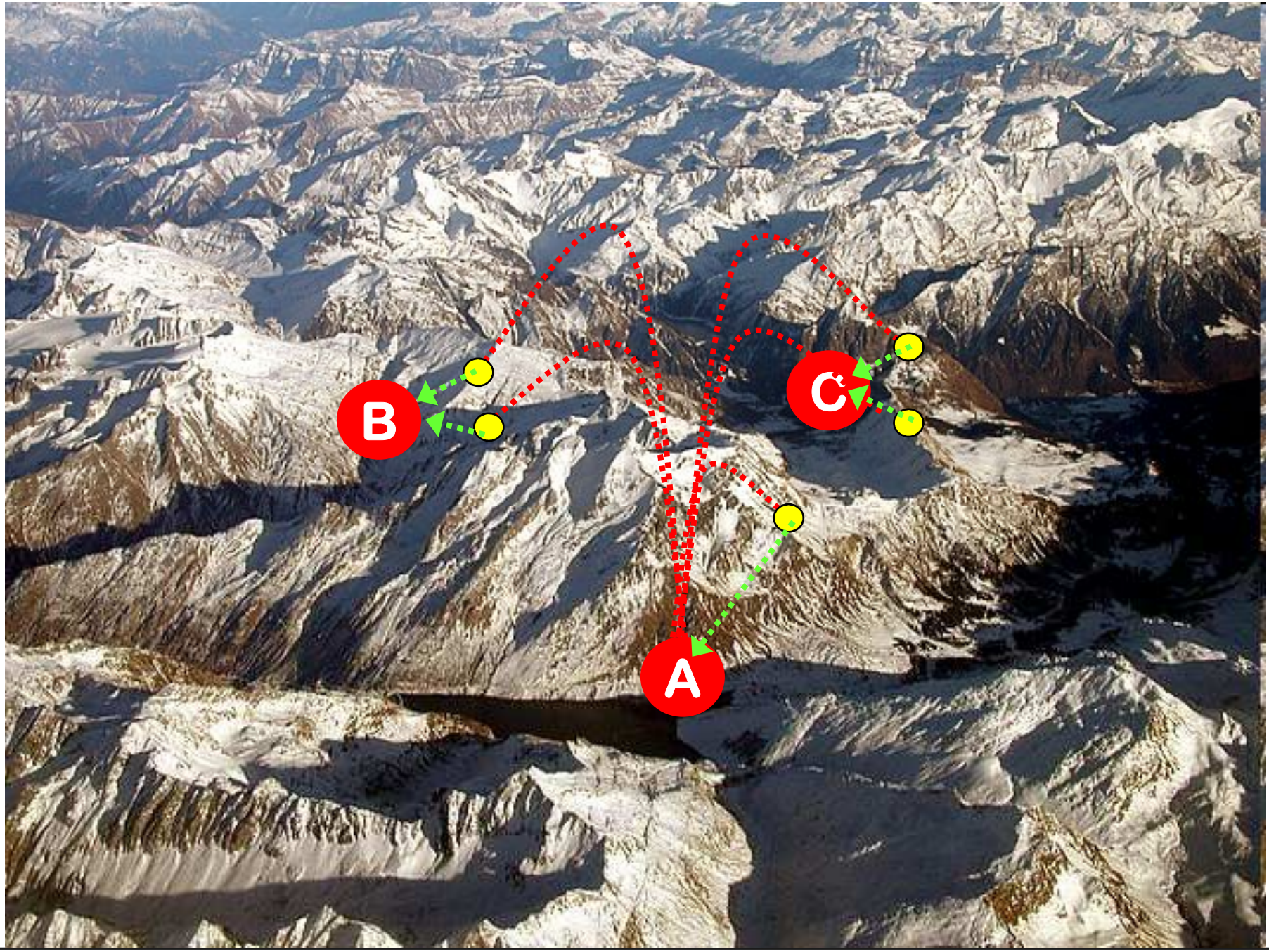


After systematic analysis, stochastic…

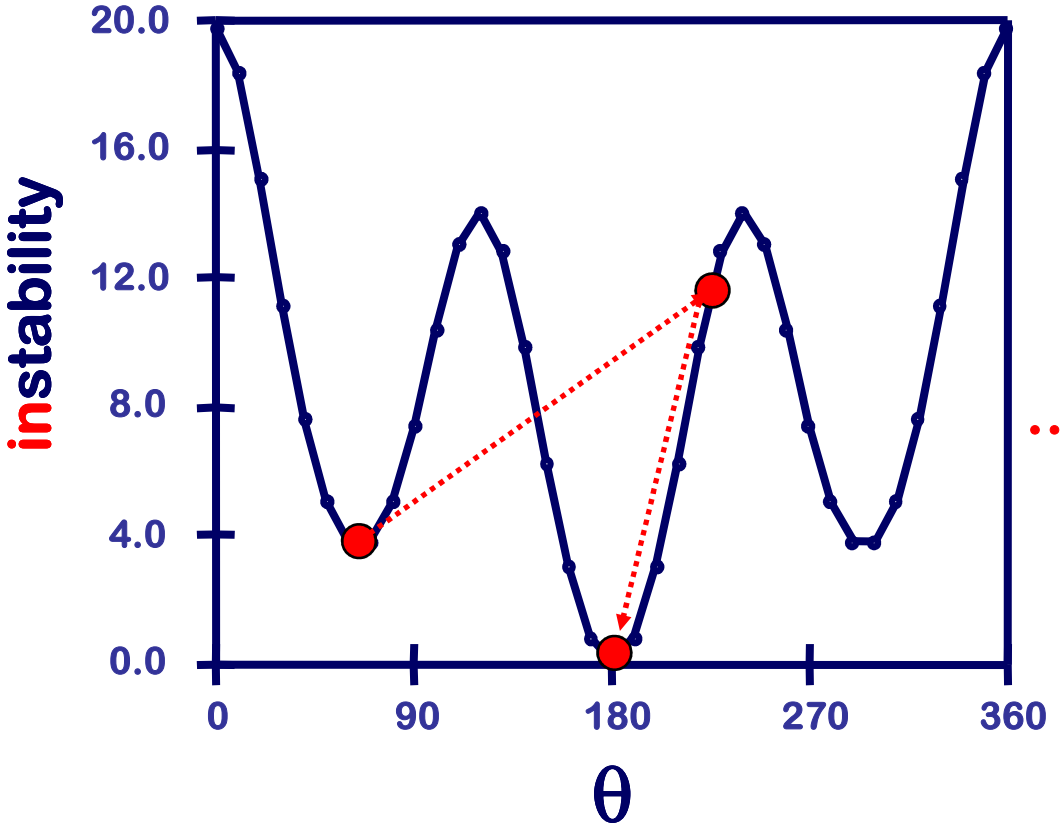


n	*θ*
0	30
1	60
2	90
…	…
12	360

n	*θ*
0	72
1	312
2	129
…	…
12	17


B
C
A

Minimization… a clear concept.



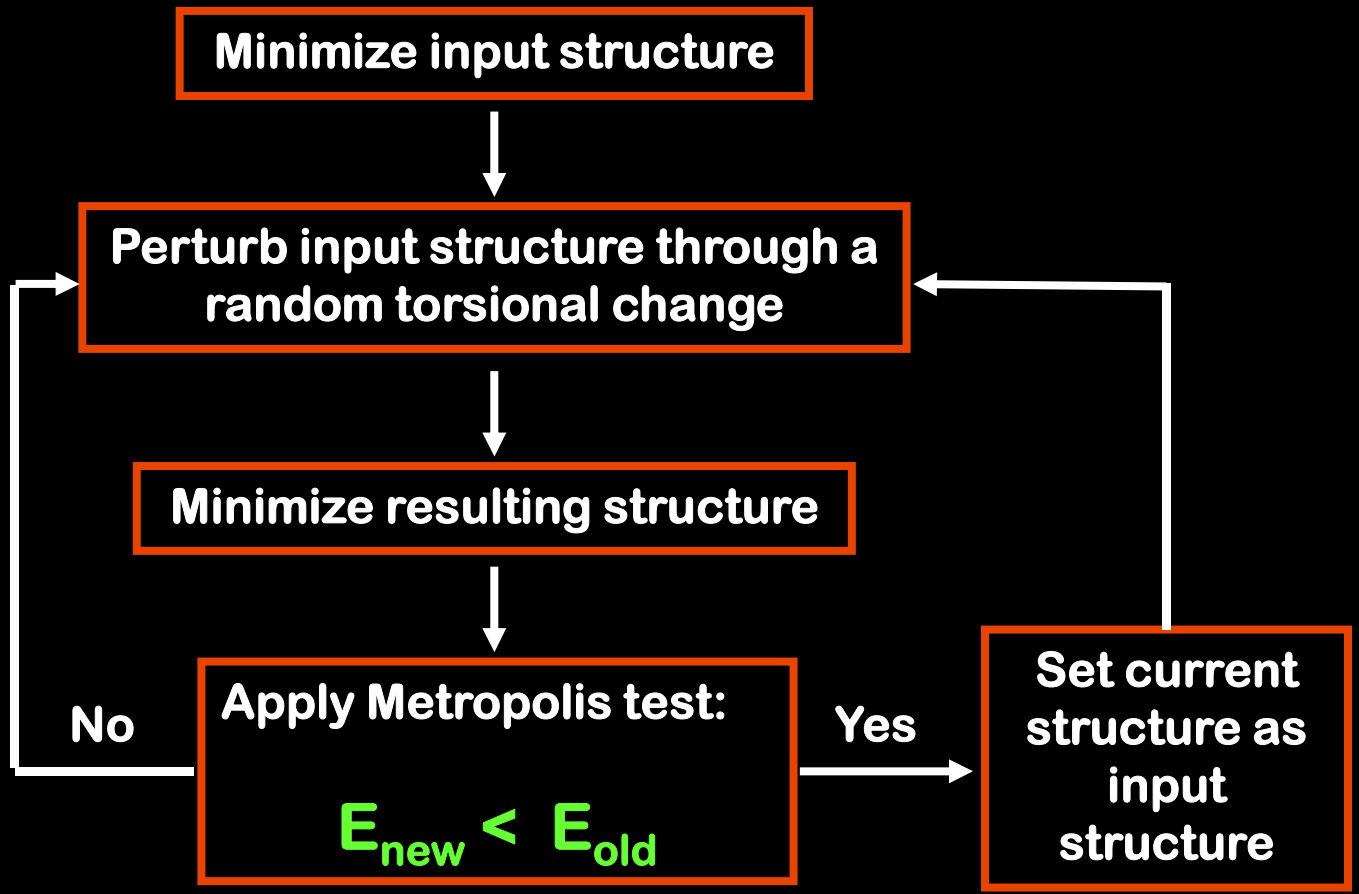
$$\frac{\Delta E}{\Delta \theta} < 0$$

… like systematic search!

$$\frac{\Delta E}{\Delta \theta} = 0$$

Monte Carlo Single Minimum

(Li & Scheraga PNAS, 84, 6611, 1987)



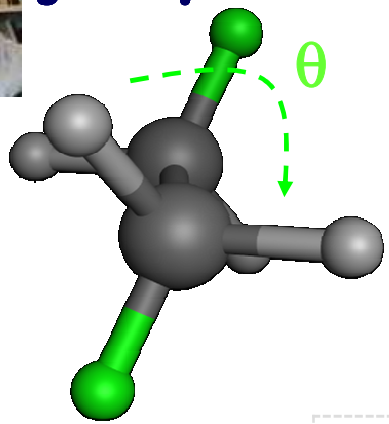
Typically, Stochastic Search is good at locating most of the local minima of flexible (*15-20 rotatable bonds*) molecules, including ring conformations and invertible tetrahedral centers (like tertiary nitrogens).

However, a large number of samples may be required to adequately search the conformation space of flexible molecules.

Stochastic Search is generally not the most efficient method for generating conformations because, like Systematic Search, it spends most of its time refining conformations that are far from the global minimum and that will ultimately be rejected.

Stochastic Search is NOT able to locate transition states… by definition!

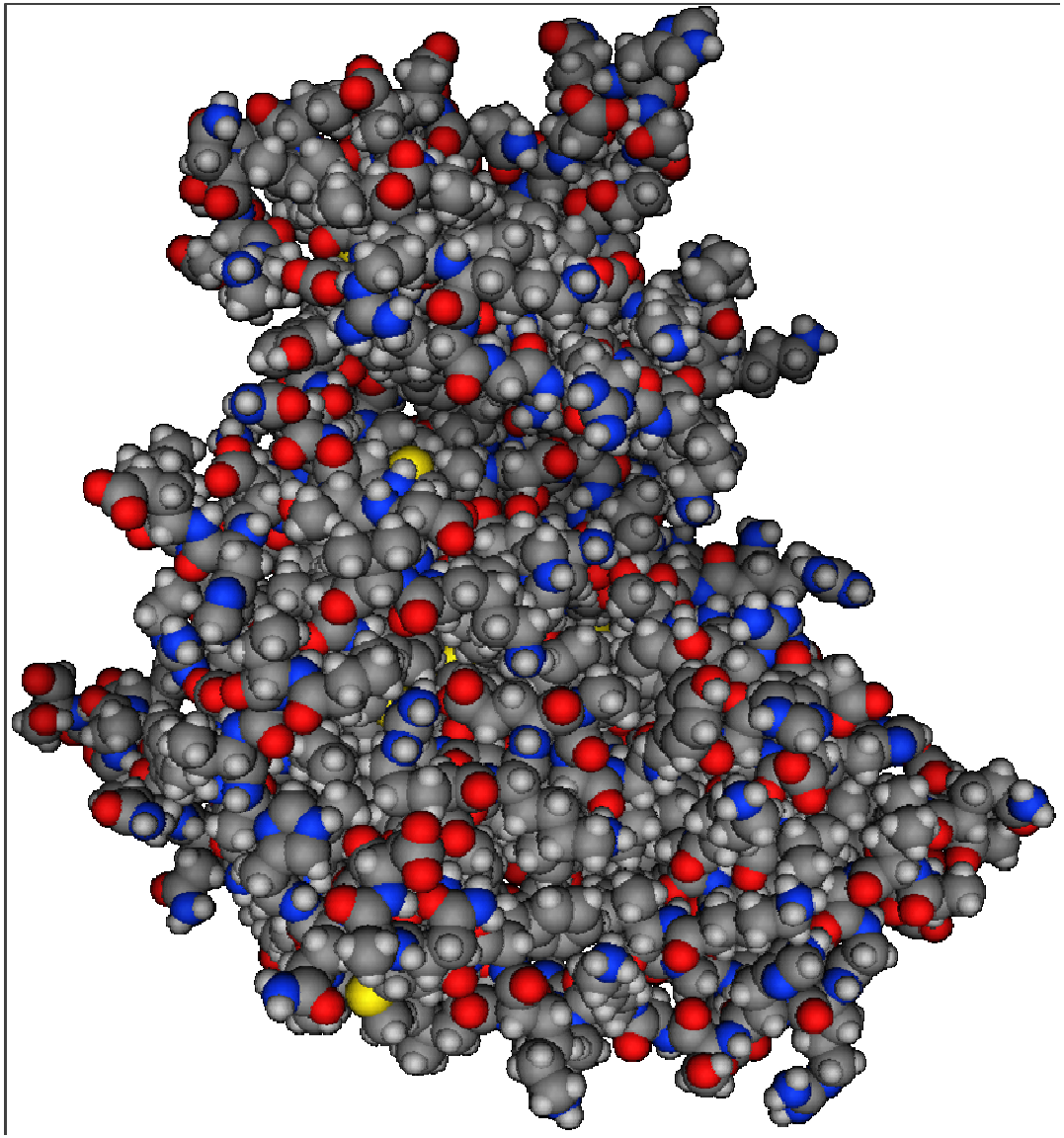
Tabu search… don't jump where you already jumped!



n	*θ*
0	72
1	312
2	129
…	…
12	17

Save θ value and skip, the next time, if the value is θ ± Δθ!!!

But we still have a very severe limitation…



And now, I would like to introduce you John Holland:

University of Michigan (USA) - 1970

… considered the father of genetic algorithms.

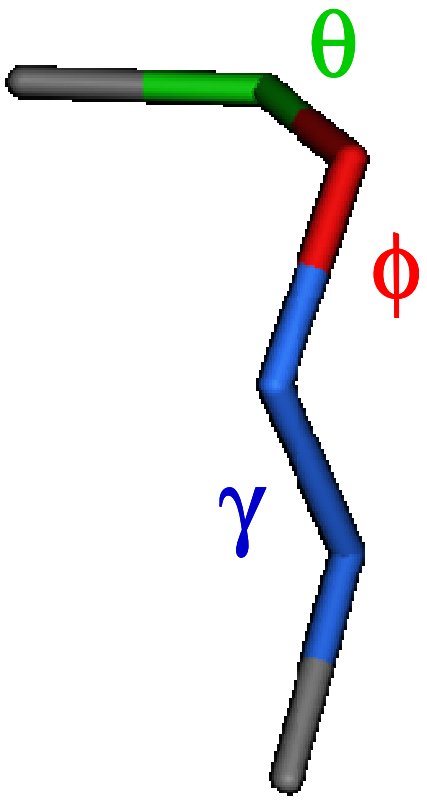
A brief overview:

Genetic Algorithms (GAs) are heuristic search algorithms premised on the evolutionary ideas of natural selection and genetic.

The basic concept of GAs is designed to simulate processes in natural system necessary for evolution, specifically those that follow the principles first laid down by Charles Darwin of survival of the fittest.

As such they represent an intelligent exploitation of a random search within a defined search space to solve a problem.

Here is the basis idea in conformer searching:



	θ	ϕ	γ
0	01001	01001	01001
30	00101	00101	00101
60	10001	10001	10001
…	…	…	…
360	01010	01010	01010

Here is the basis idea in conformer searching:

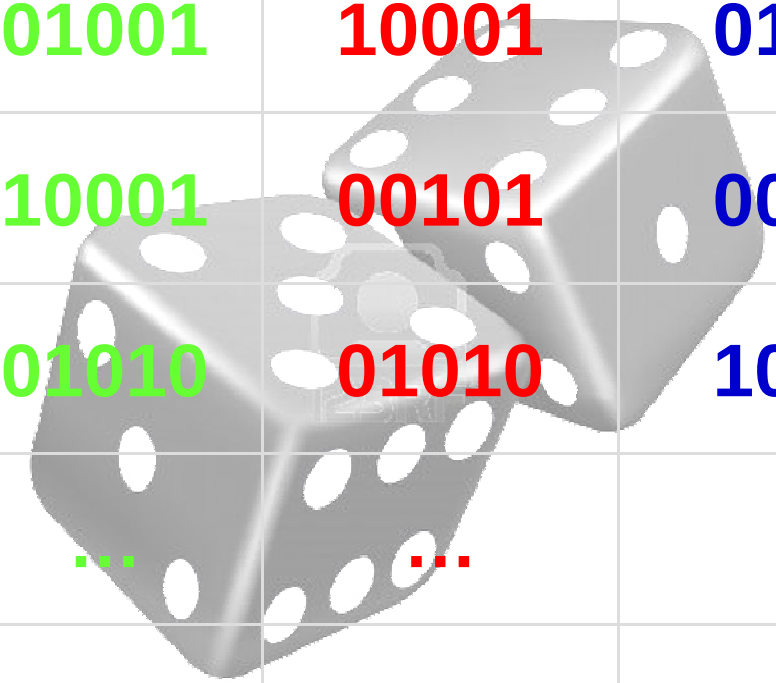
	θ	ϕ	γ
0	01001	01001	01001
30	00101	00101	00101
60	10001	10001	10001
…	…	…	…
360	01010	01010	01010

GENE

CHROMOSOME

STEP 1: randomly generate an initial population… of conformers

CHROMOSOME	θ	ϕ	γ
1	01001	10001	01010
2	10001	00101	00101
3	01010	01010	10001
…	…	…	…
n	01010	01010	01001



Evolutionary operators: the rules of evolution.

1. REPRODUCTION (selection probability);
2. RECOMBINATION (cross-over);
3. MUTATION

STEP 2: compute and save the *fitness* for each individual in the current population

CHROMOSOME	θ	ϕ	γ	FITNESS
1	01001	10001	01010	E_1
2	10001	00101	00101	E_2
3	01010	01010	10001	E_3
…	…	…	…	…
n	01010	01010	01001	E_4

STEP 3: define selection probabilities for each individual proportional to fitness

Initial population

Chromosome 1:
01001 00101 10001 01010

Chromosome 2:
01001 00101 10001 11011

Chromosome 3:
01001 00101 10001 01011

FITNESS

E_1 = 5 kcalmol

E_2 = -30 kcalmol

E_3 = -15 kcalmol

REPRODUCTION OPERATOR

New generation

Chromosome 1: x 1
01001 00101 10001 01010

Chromosome 2: x 5
01001 00101 10001 11011

Chromosome 2:
01001 00101 10001 11011

Chromosome 2:
01001 00101 10001 11011

Chromosome 2:
01001 00101 10001 11011

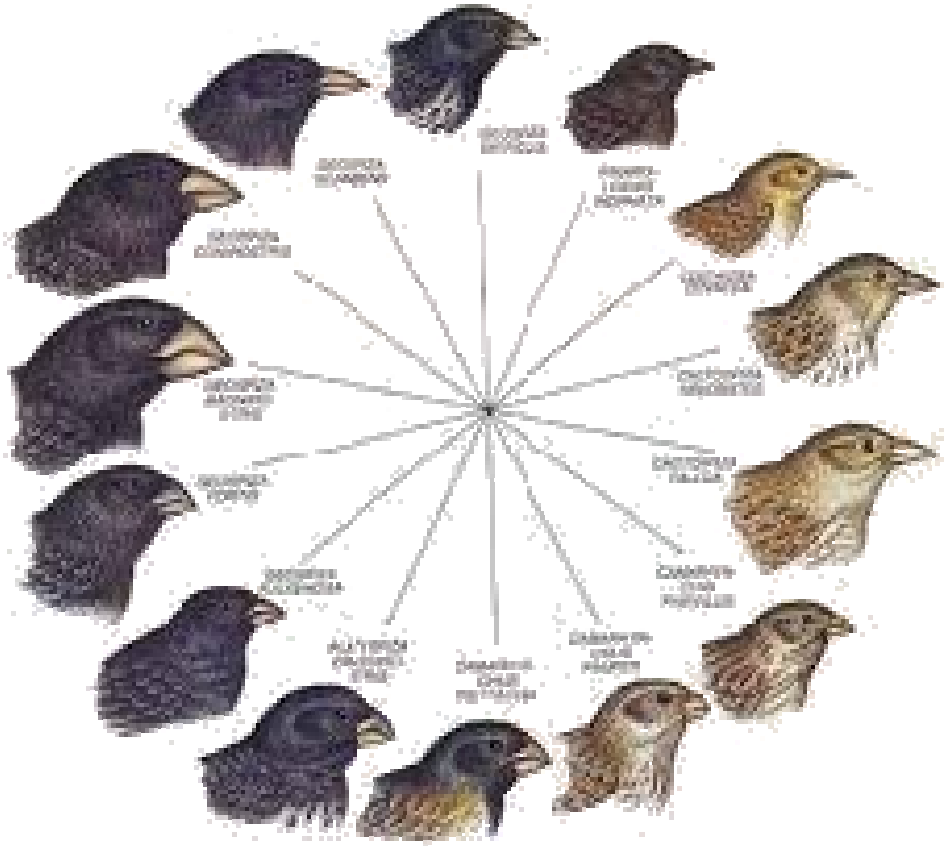
Chromosome 2:
01001 00101 10001 11011

Chromosome 3: x 3
01001 00101 10001 01011

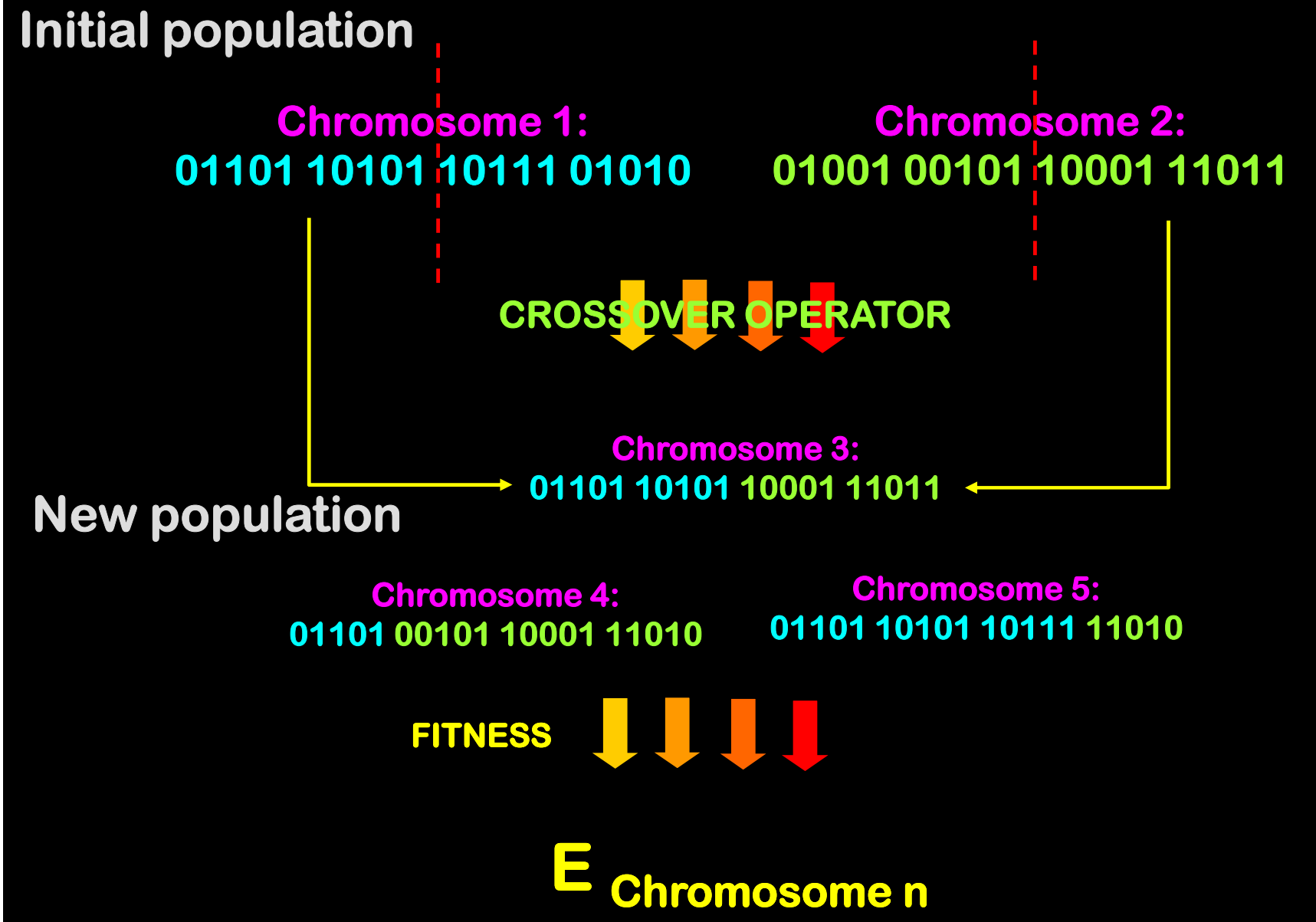
Chromosome 3:
01001 00101 10001 01011

Chromosome 3:
01001 00101 10001 01011

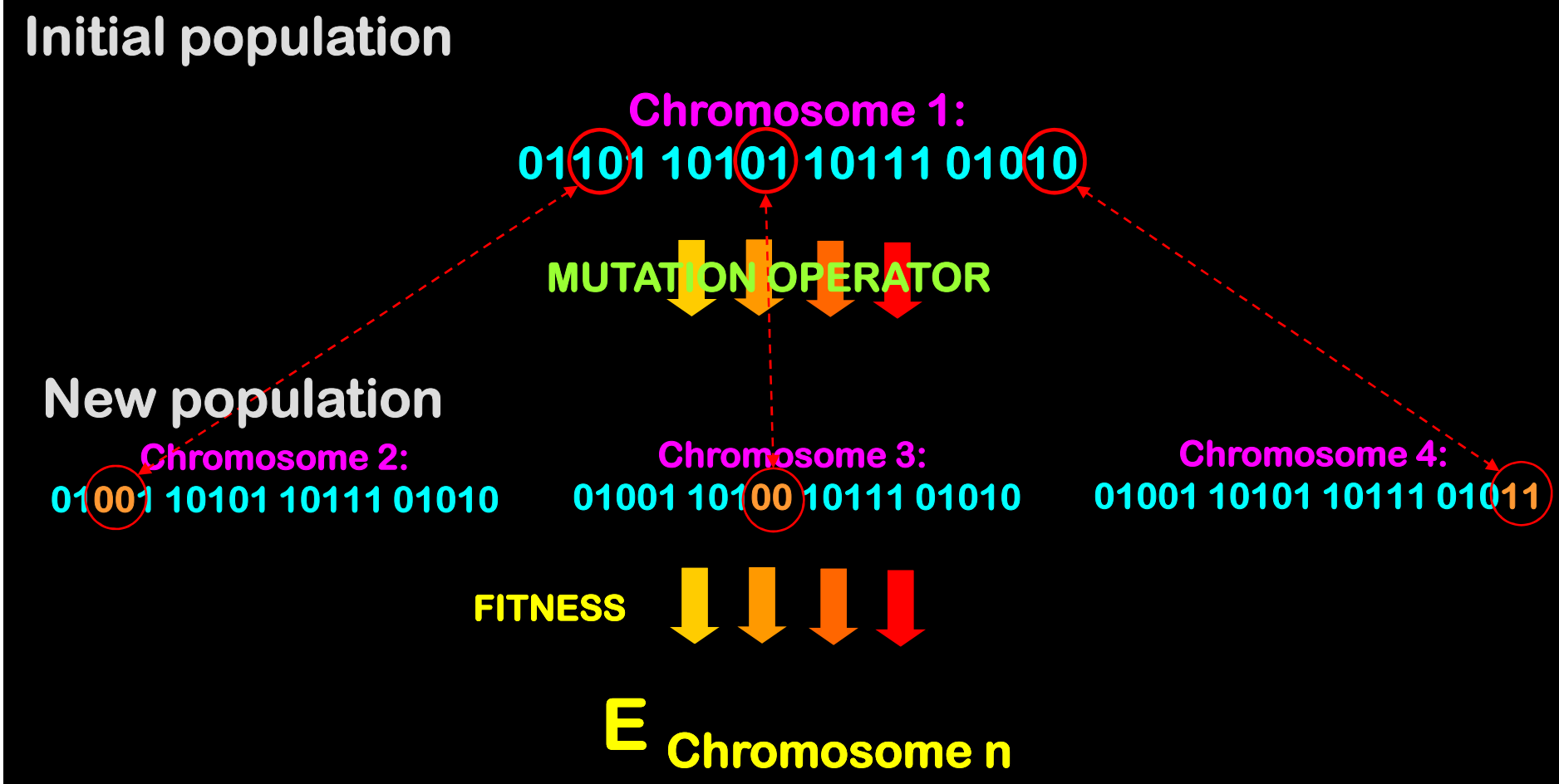
STEP 4: Generate new generations by probabilistically selecting individuals from the initial one to produce offspring via genetic operators



CROSS-OVER OPERATOR



MUTATION OPERATOR

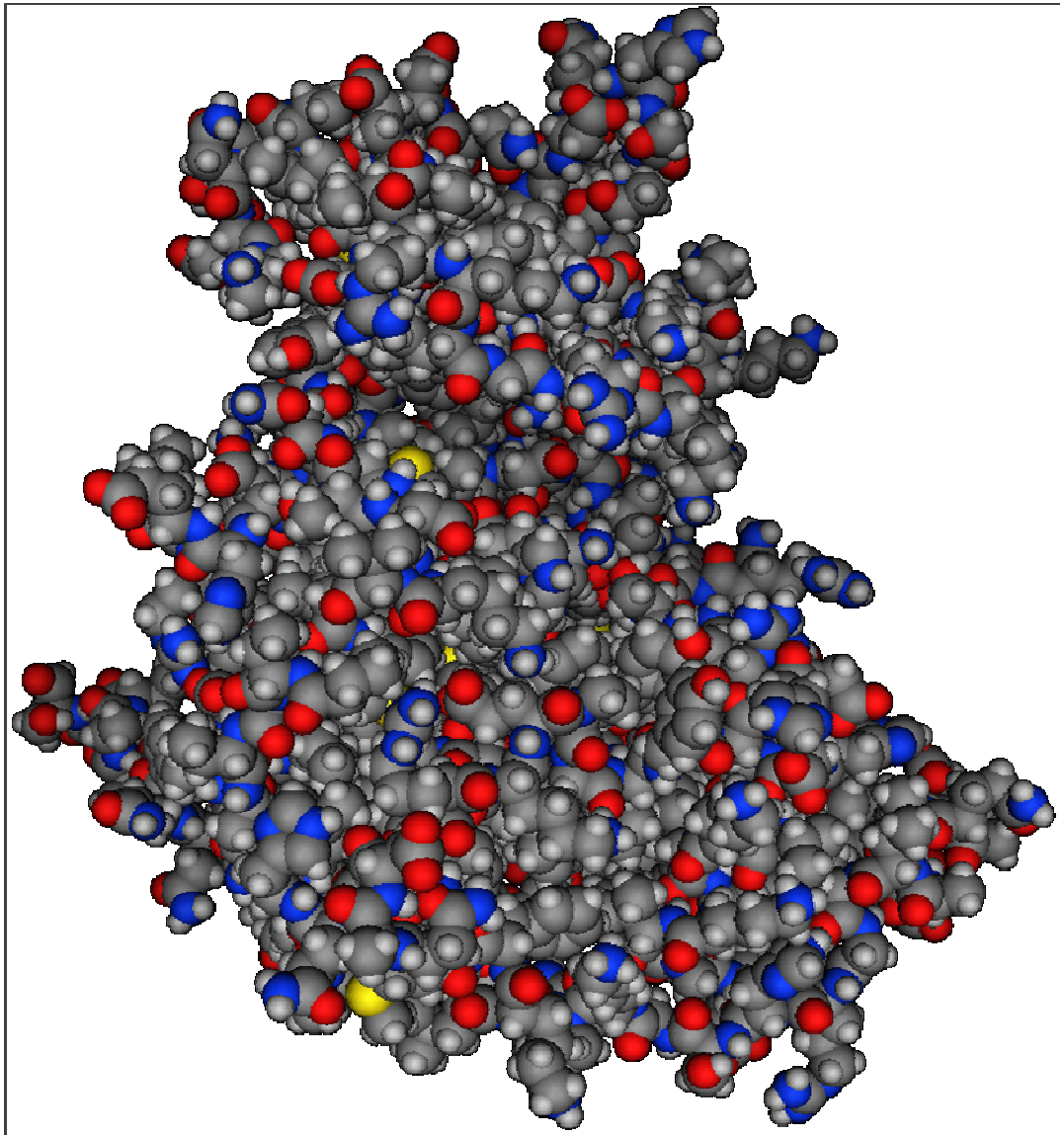


NB: mutation operator is usually applied with low frequency and often in the early stage of evolution of the initial population. In fact, it is easily understandable that mutation operator can drastically change the fitness of the individual during its evolution.

STEP 5: repeat step 2 until satisfying solution is obtained.

Though it might not find the best solution. More often than not, it would come up with a partially optimal solution.

But we still have a very severe limitation…




TITLE
"Progettazione e Sviluppo di un Farmaco"
DIRECTOR
Stefano Moro
CAMERA
Chimica e Tecnologia Farmaceutiche
DATE
SCENE
TAKE