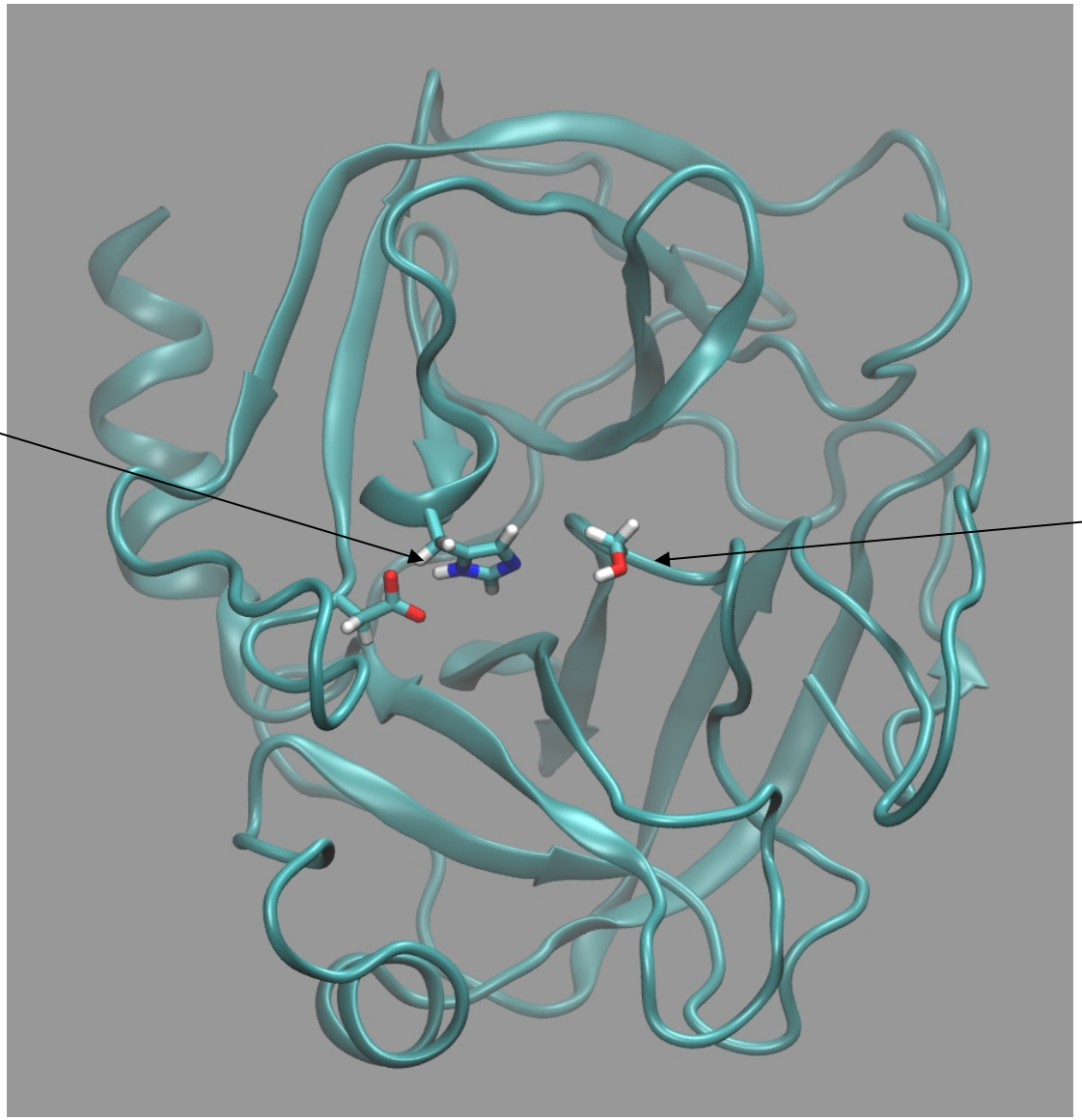
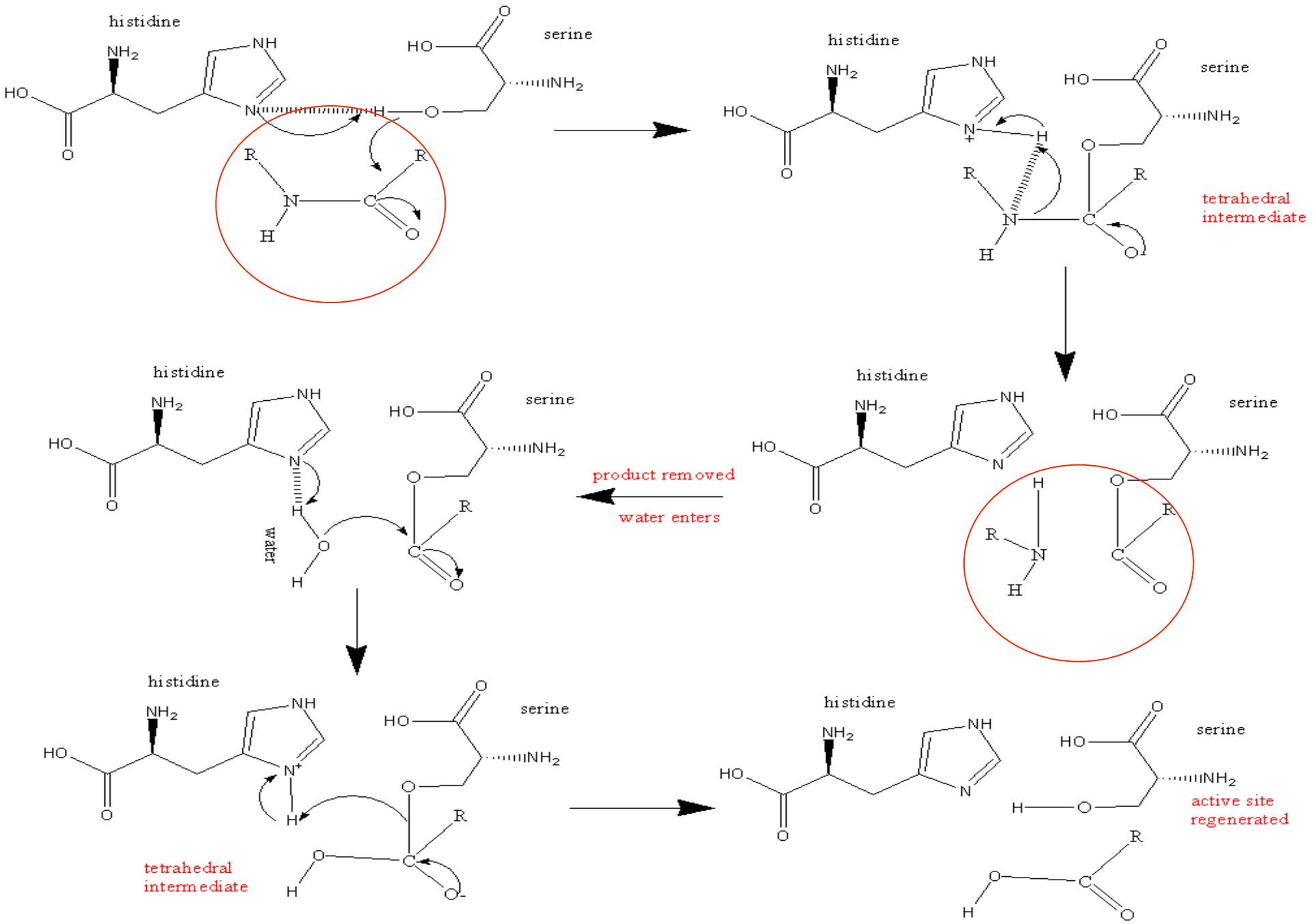
The Glass Ceiling of Biomolecular Computation

Trevor Gokey
San Francisco State University
Department of Computer Science

Serine Proteases – Trypsin below

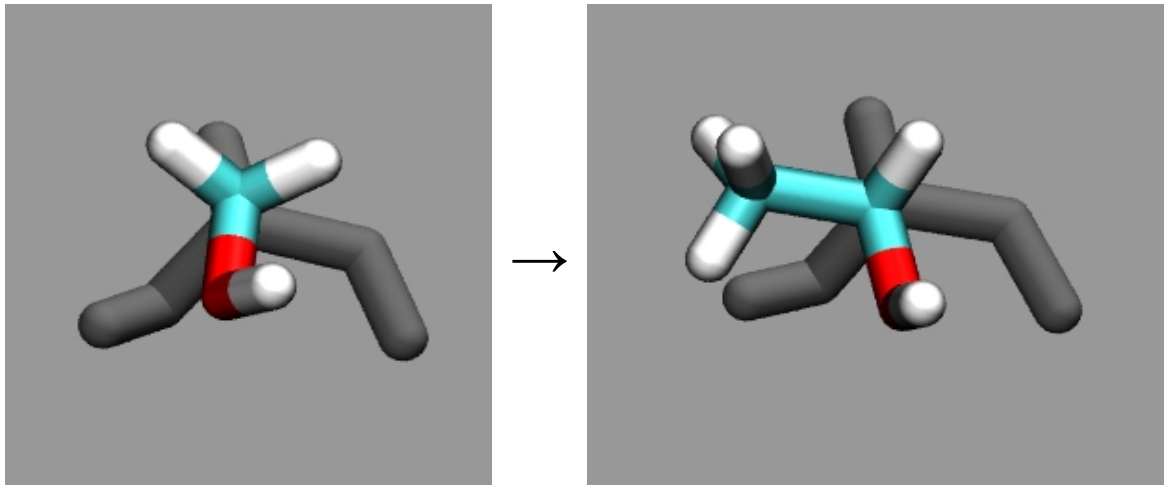


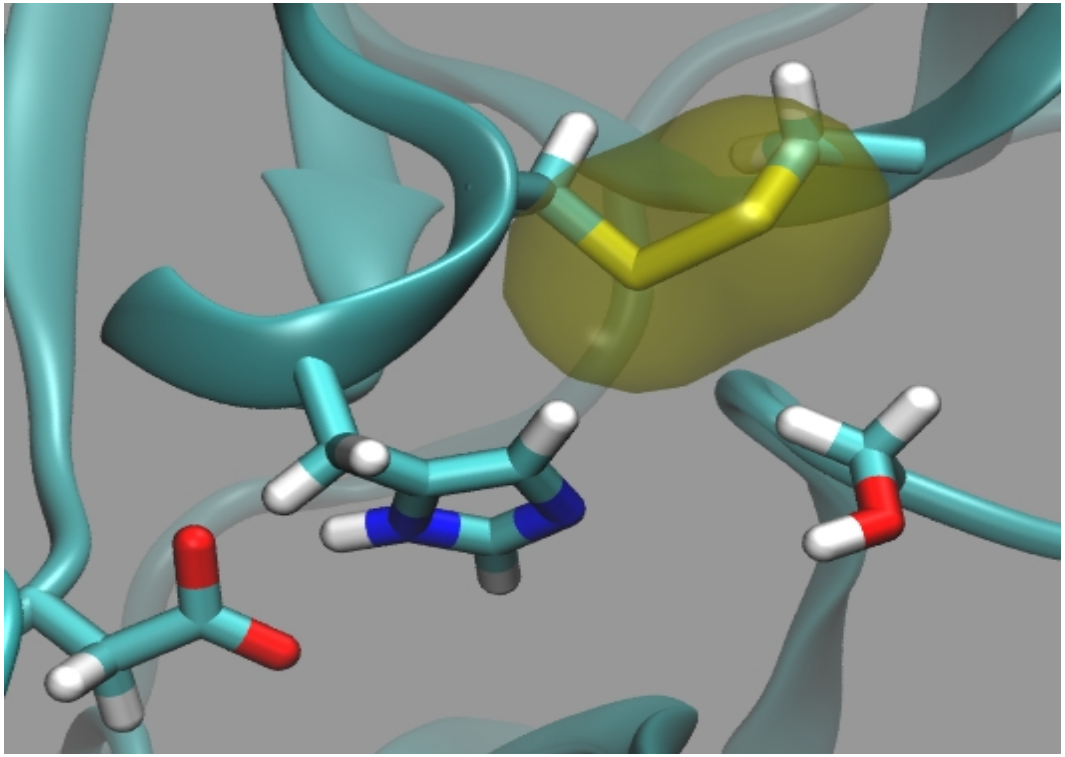
Highly conserved protein--useful for digestion and stain removal



http://en.wikipedia.org/wiki/File:Serine_protease_mechanism_by_snellios.png

serine → threonine



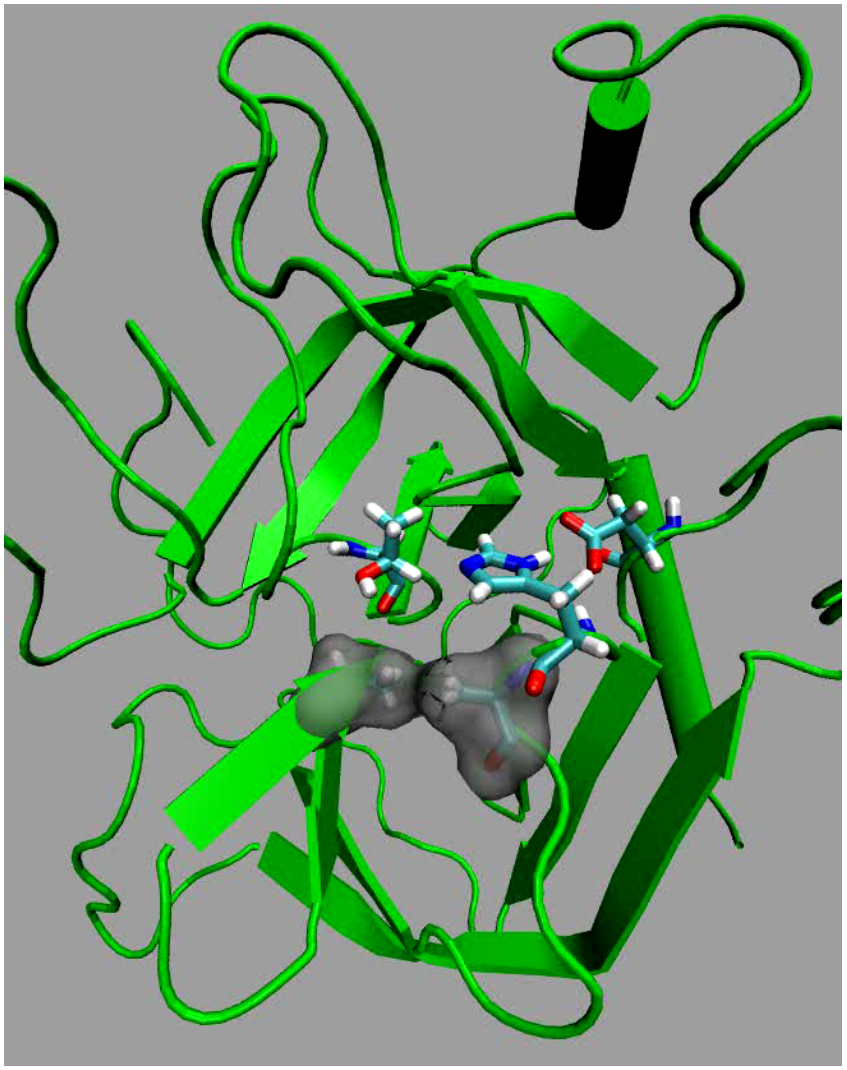


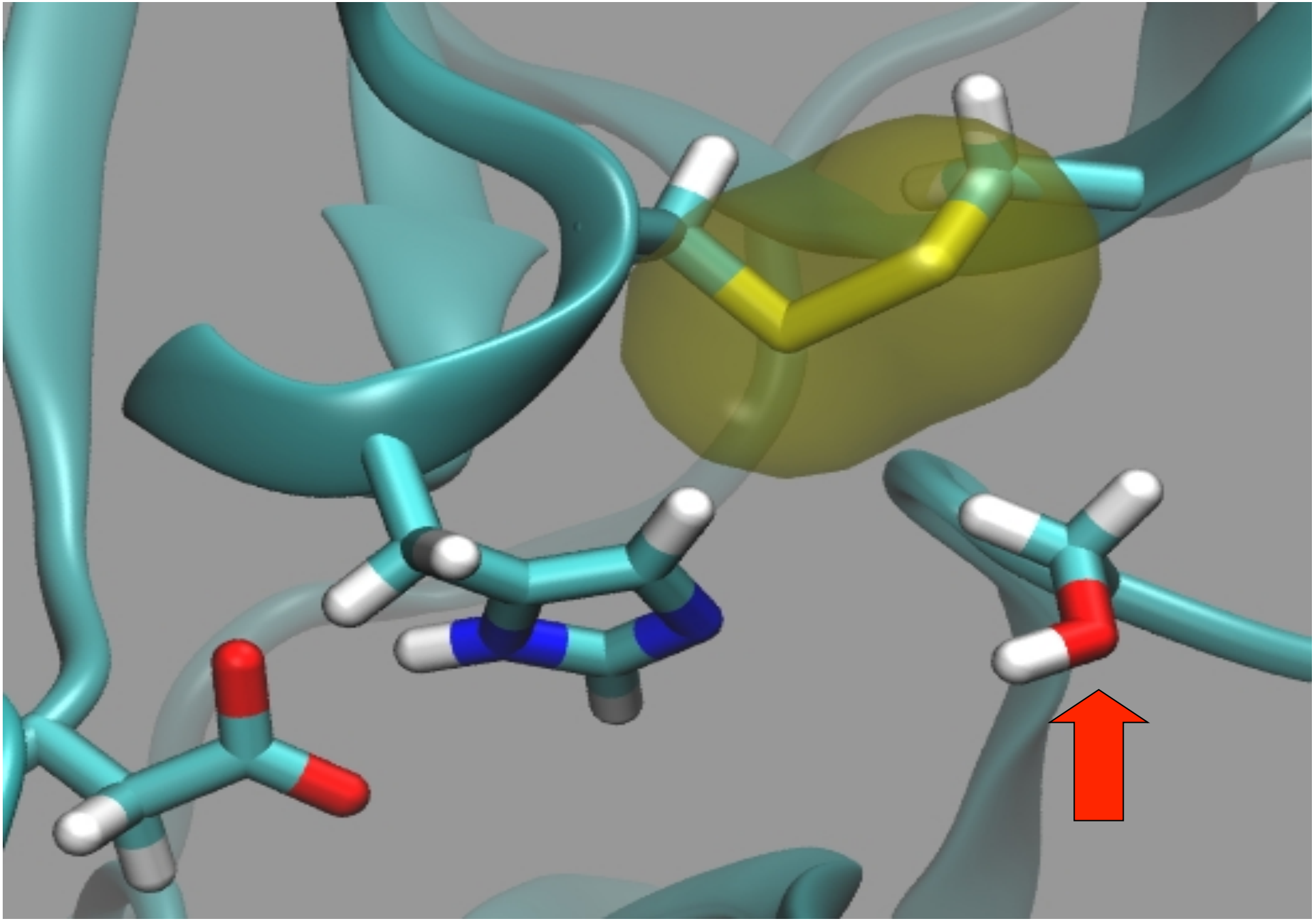
Wet lab:

Disulfide bridge is bulky

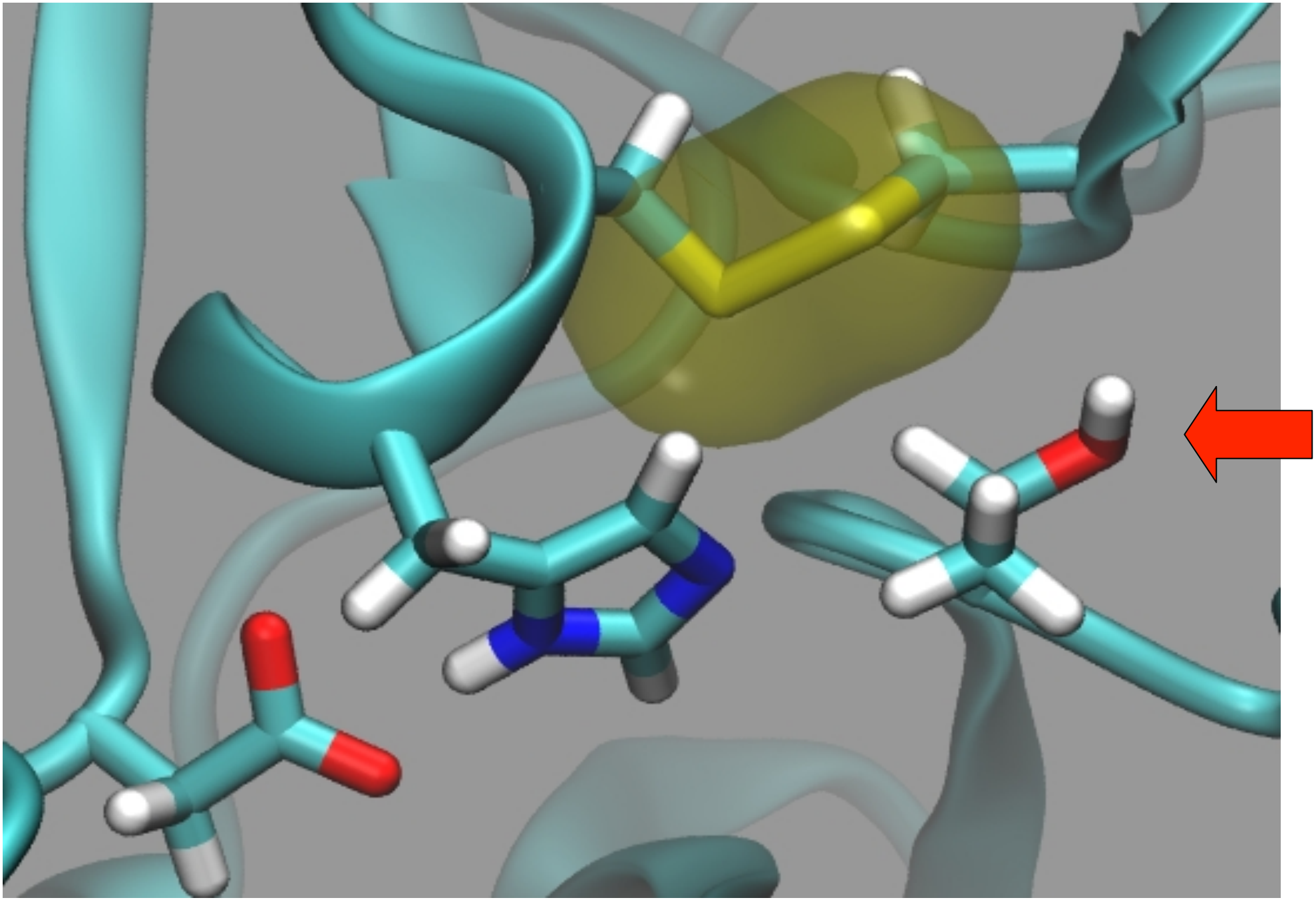
→ Modify for threonine to fit

Bring out the “computational microscope”

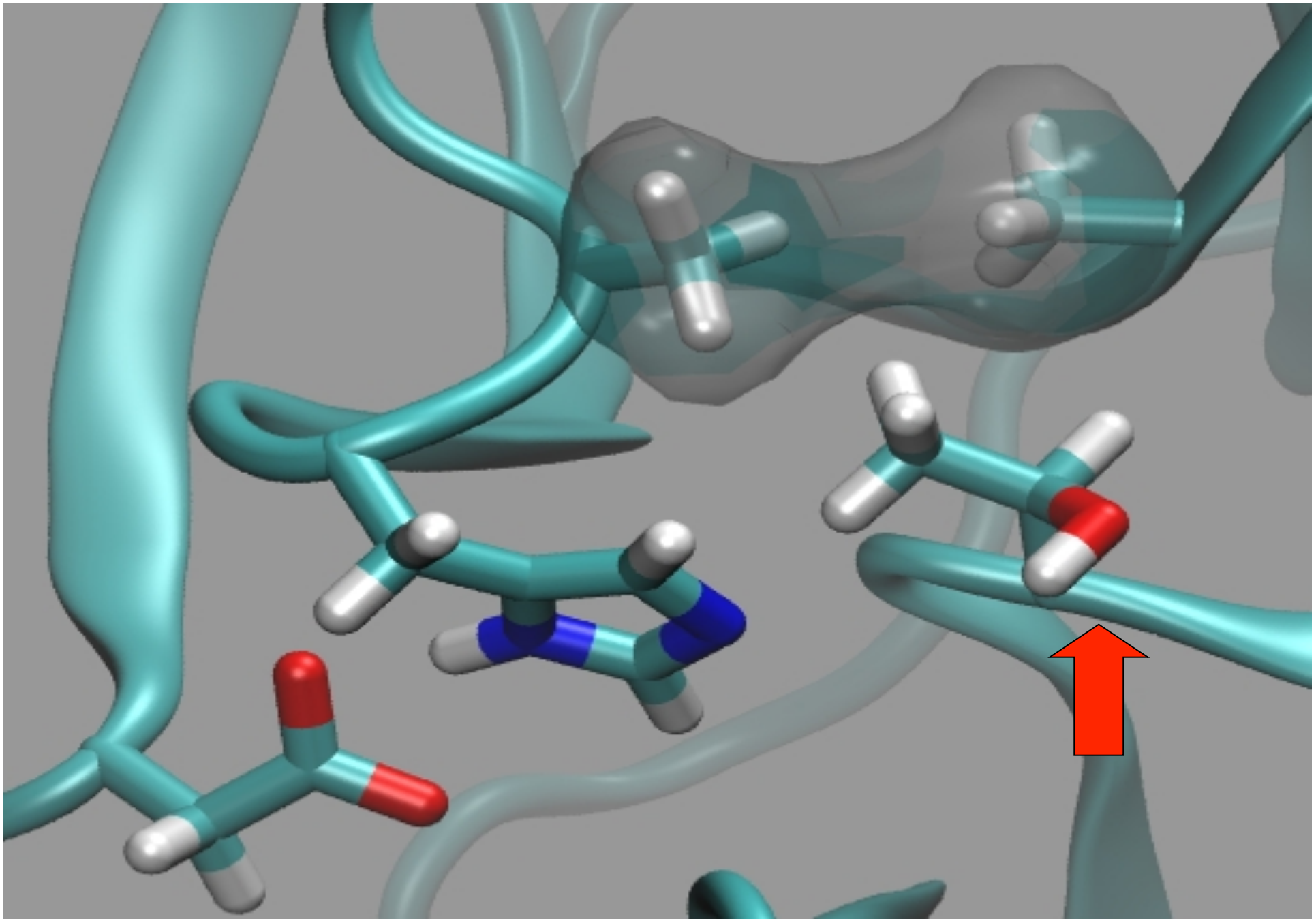




Active



Inactive

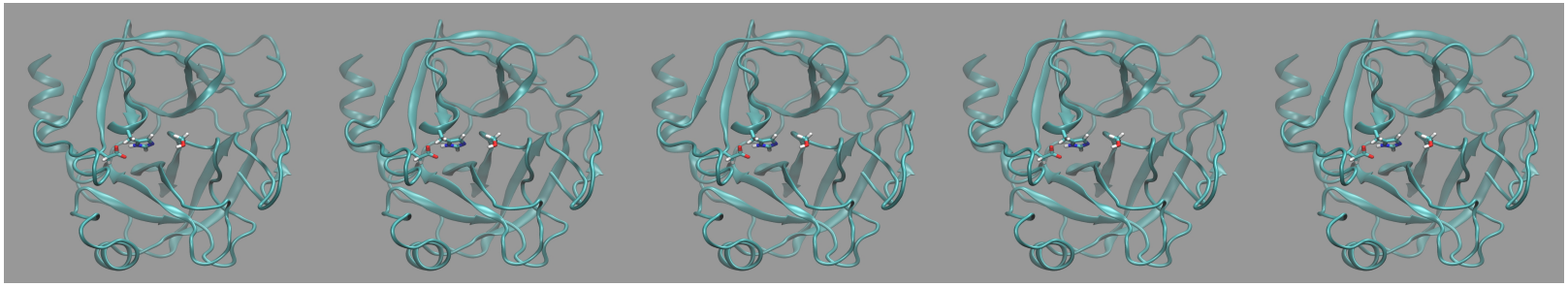


Reduced activity

5 proteins • 115 ns = 575 ns

1 year ago → 12 ns•day^{-1}

2 month wall time

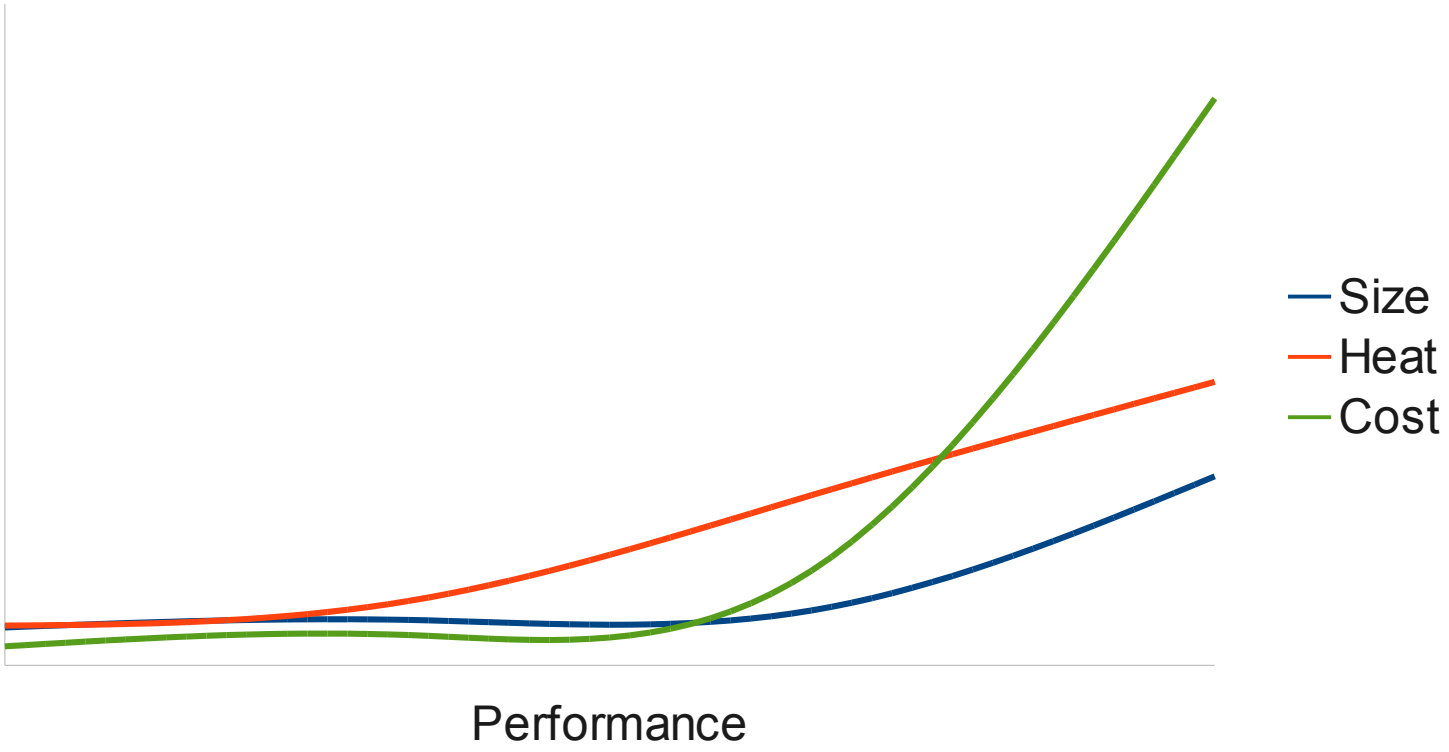


June → **August**

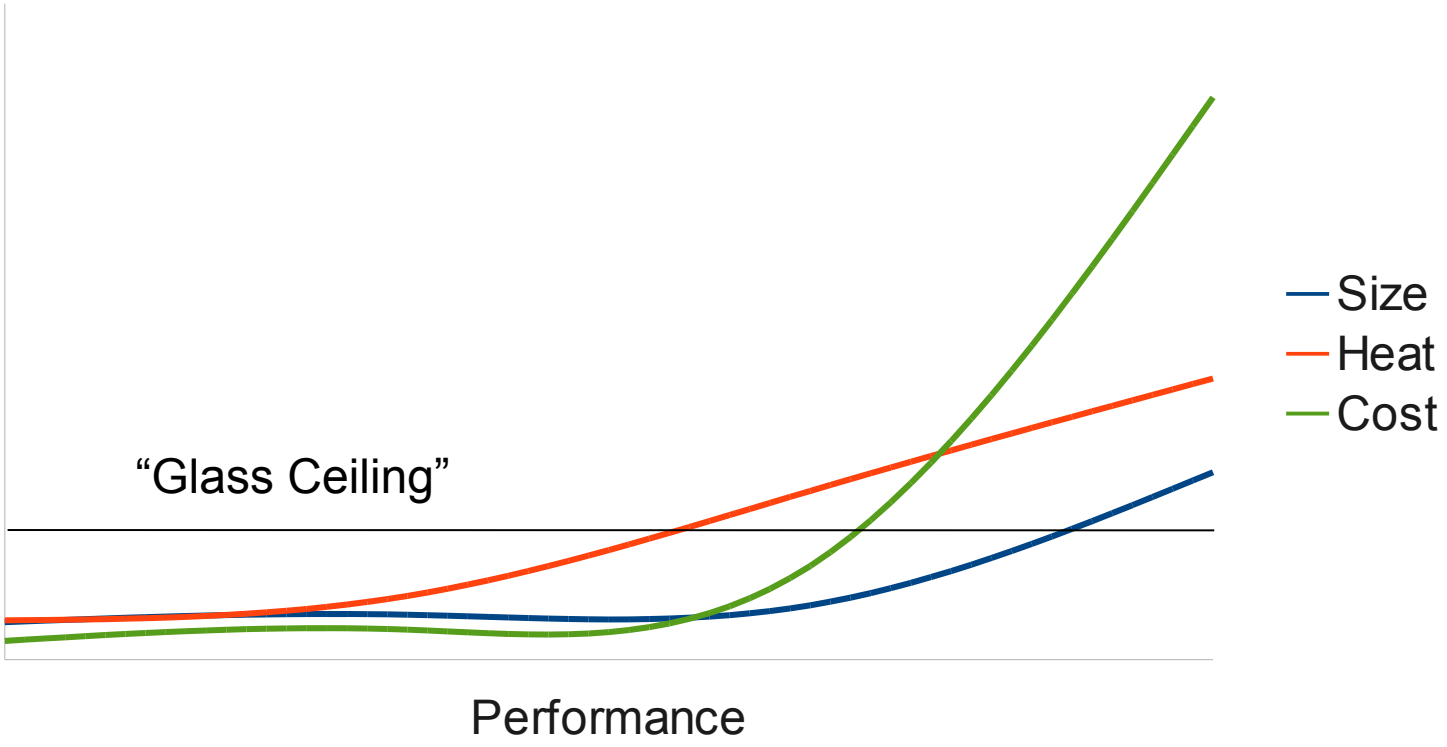
Maximize: performance

Minimize: size, heat, cost

Where is the *sweet spot*?



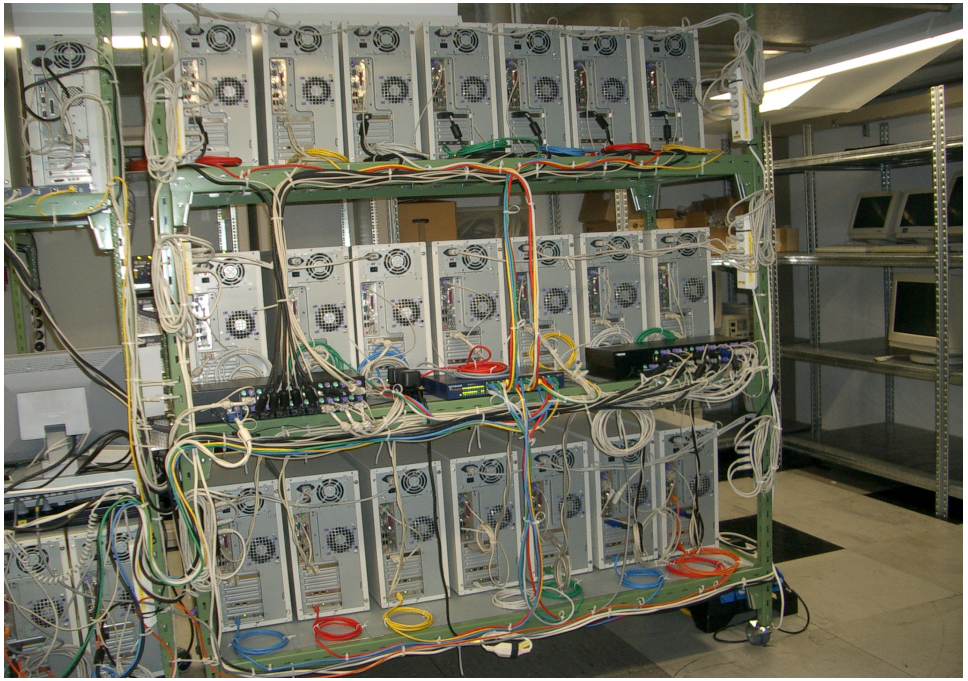
Hardware for small labs



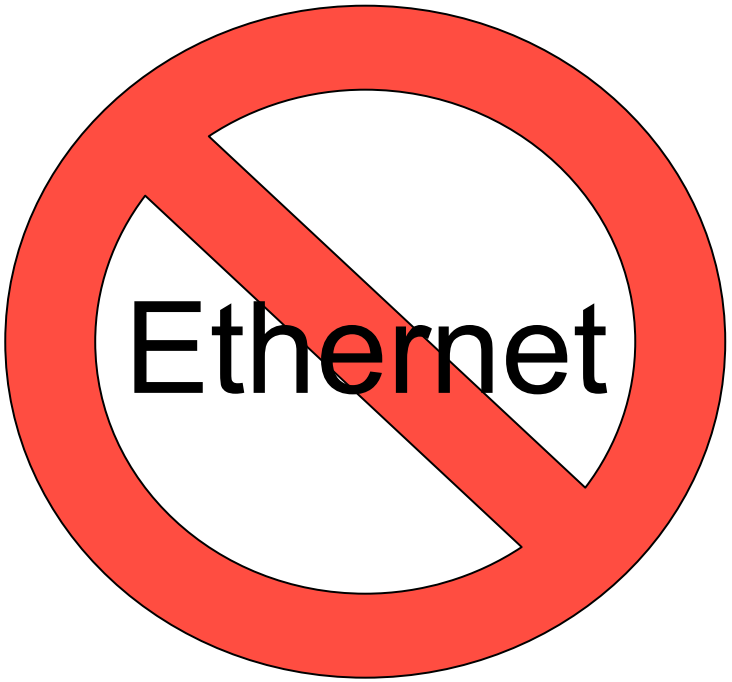
1 Gb/s Ethernet:

Low cost

Low performance



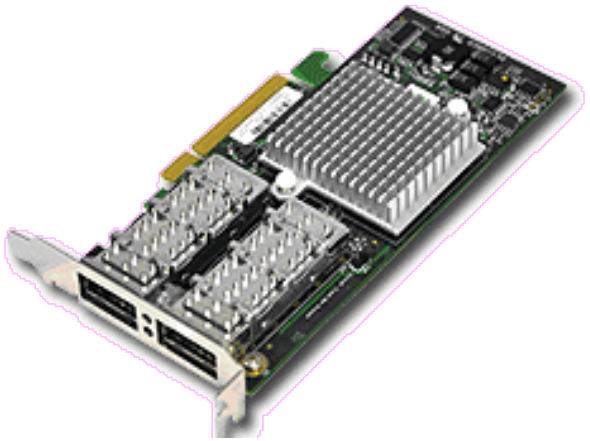
http://www.tik.ee.ethz.ch/~ddosvax/cluster/img/photos/cluster_back.jpg

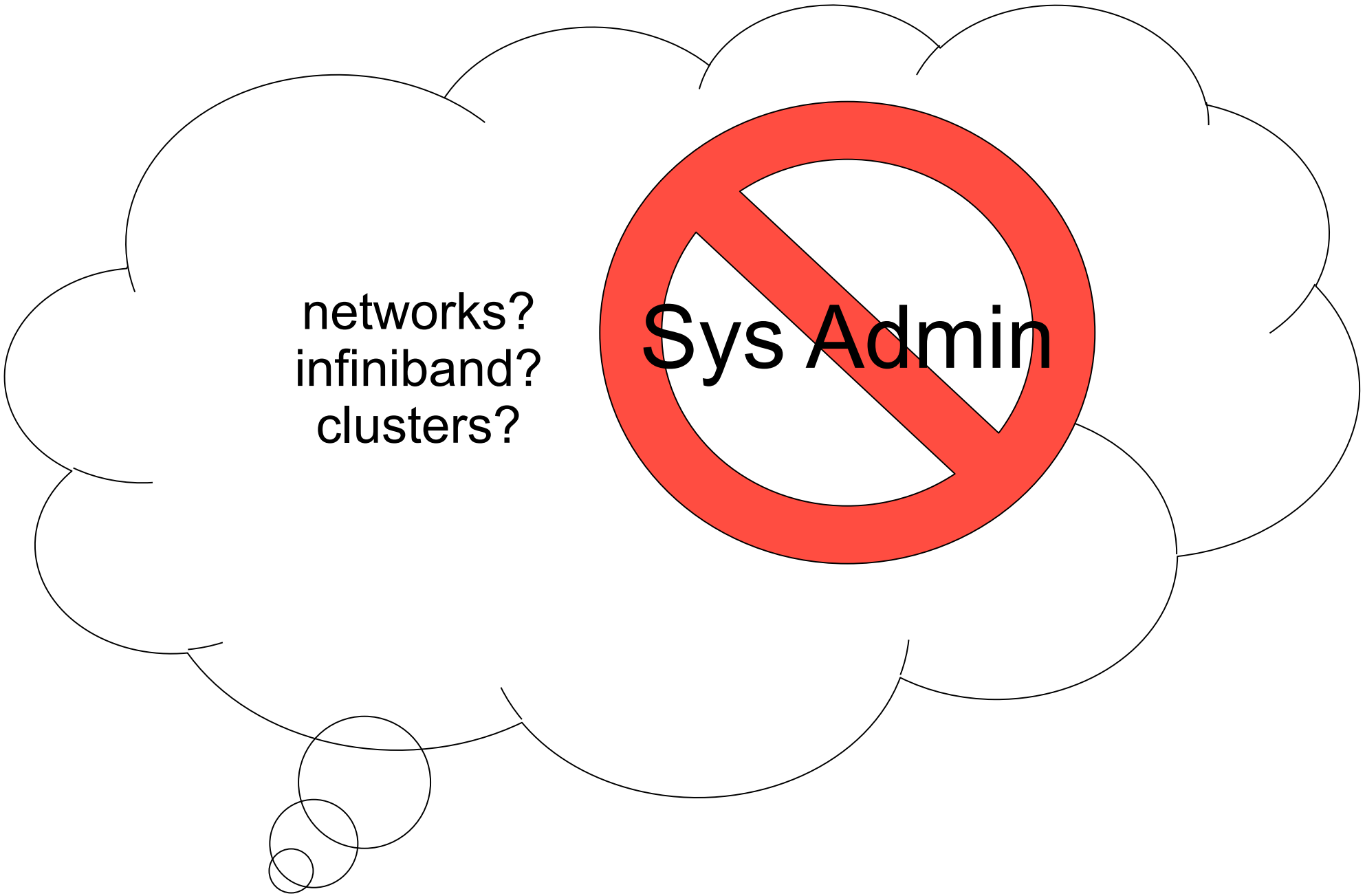


Infiniband:

Medium cost

Medium performance

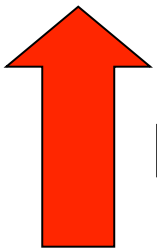




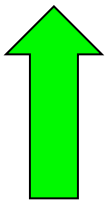
Chemist

CPU or GPU?

32 core machine:

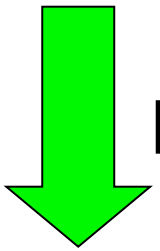


High cost

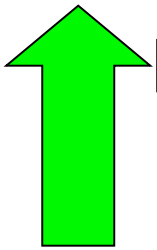


Medium performance

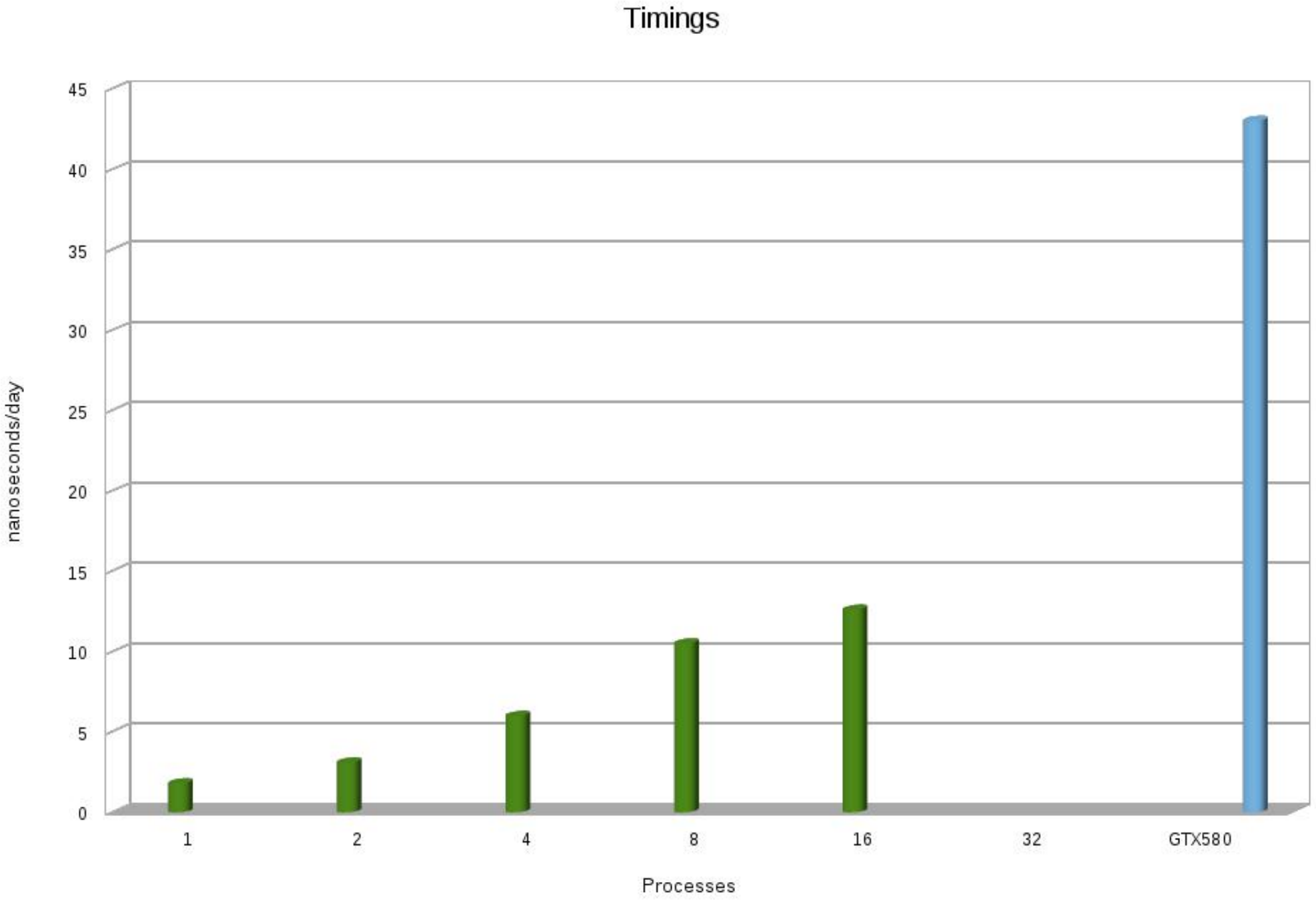
GTX580:

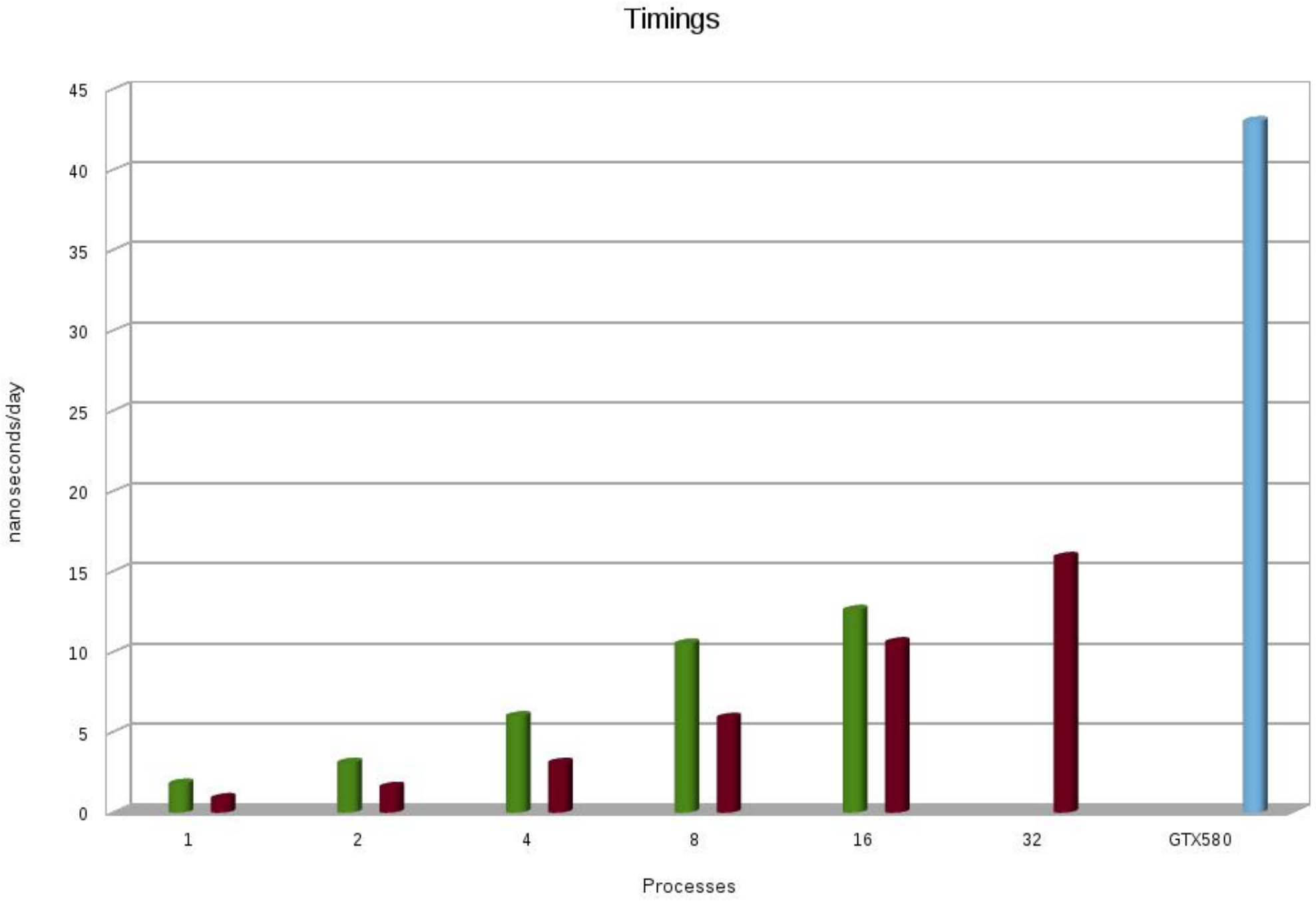


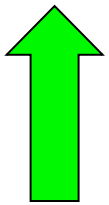
Low cost

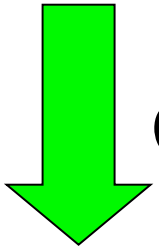


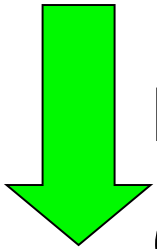
High performance


Timings
45
40
35
30
25
20
15
10
5
0
nanoseconds/day
1
2
4
8
16
32
GTX580
Processes


Timings
45
40
35
30
25
20
15
10
5
0
nanoseconds/day
1
2
4
8
16
32
GTX580
Processes





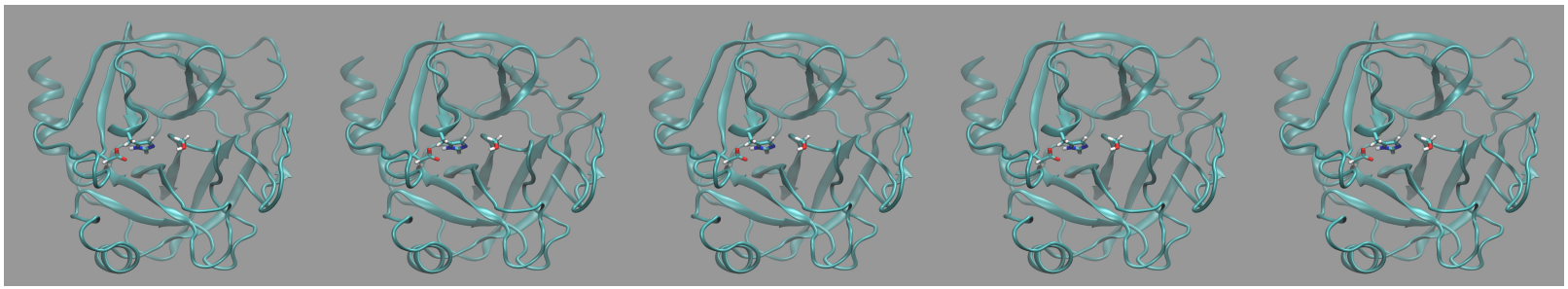


	GPU		CPU	
Speed ratio:	43 ns•day^{-1}	:	15 ns•day^{-1}	= 2.87
Cost ratio:	$~500	:	$~2000	= 0.25
Power ratio: (mfg. rating)	244 W	:	460 W	= 0.53

5 proteins • 115 ns = 575 ns

Now → 43 ns•day^{-1}

2 months → 2 weeks



December → **December**

Results over winter break

Looking forward: Utilize MPI CUDA capabilities of AMBER