True landscape of drug-protein binding

T.Matsuzaki*, T.Tanaka, K.Ikeda and H.Komatsu

* tak-matsuzaki-45@cb3.so-net.ne.jp

Interprotein Corporation

Abstract

"Docking" is a major tool in *in-silico* drug design and the method has been extensively developed. Simultaneously, computer hardware has been developed quite a lot. Nevertheless, poor performance of the docking methods are well known and its reason is not explained so far.

We have tried to understand the situation and reached a plausible explanation. The direct output of MO or MD calculation for a binding energy between drug and protein, is in a range of 40-60 kcal/mol. We subtract a constant to make calculated value comparable with experimentally observed value, 8-12 kcal/mol. On the other hand, drug activity measure, Kd, is defined as kT ln Kd = DG. And $\Delta G = \Delta H - T\Delta S$. MO and MD calculate only ΔH term and their output of binding energy either neglects $T\Delta S$ term or assumes $T\Delta S$ is similar for all compounds. The reason of the poor performance of docking methods seems to lie in this process.

Contribution of entropy must be significant if we consider released solvation water upon binding and restricted freedom of bound small molecules. We assume the observed binding energy of -9 kcal/mol must be the sum of -6 kcal/mol and -3 kcal/mol for ΔH and $-T\Delta S$, respectively. Although you can not judge 1 μM activity based on calculated binding energy in a range over -6 kcal/mol, MO and MD programs output -9 kcal/mol after artificial adjustment. Consequently, the selected compounds by docking methods do not show activity in the most cases.

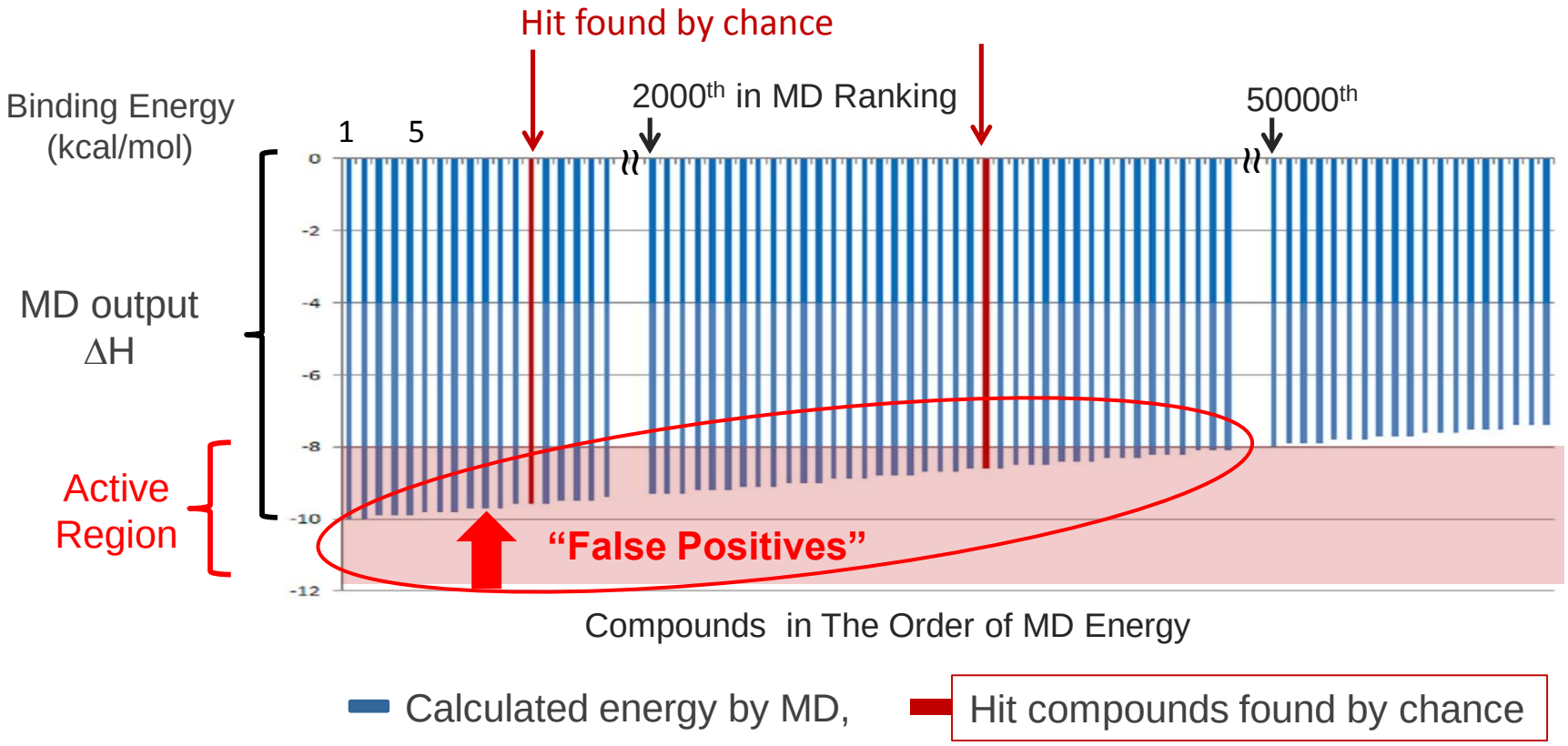
Interprotein has developed a system which consider entropy contribution and concentrates compounds with high total energy into higher rank. It has been successfully applied for 20 PPI targets and 10 enzyme targets. Details in Runx1-CBFβ target is shown.

Conclusion

- The reason why elaborate calculations on best supercomputers do not succeed in finding active compounds is explained by the lack of considerations on “Entropy”.
- We assume correct active binding structures exist among top 50,000 computer output, but huge amount of false positives hide correct structures.
- Consideration on entropy will greatly help to find active structures and our INTENDD ® system have continuous success on 10 enzymes and 20 PPI targets.
- Example on RUNX1-CBFb target is shown in detail.

Docking problem | High false positive rate

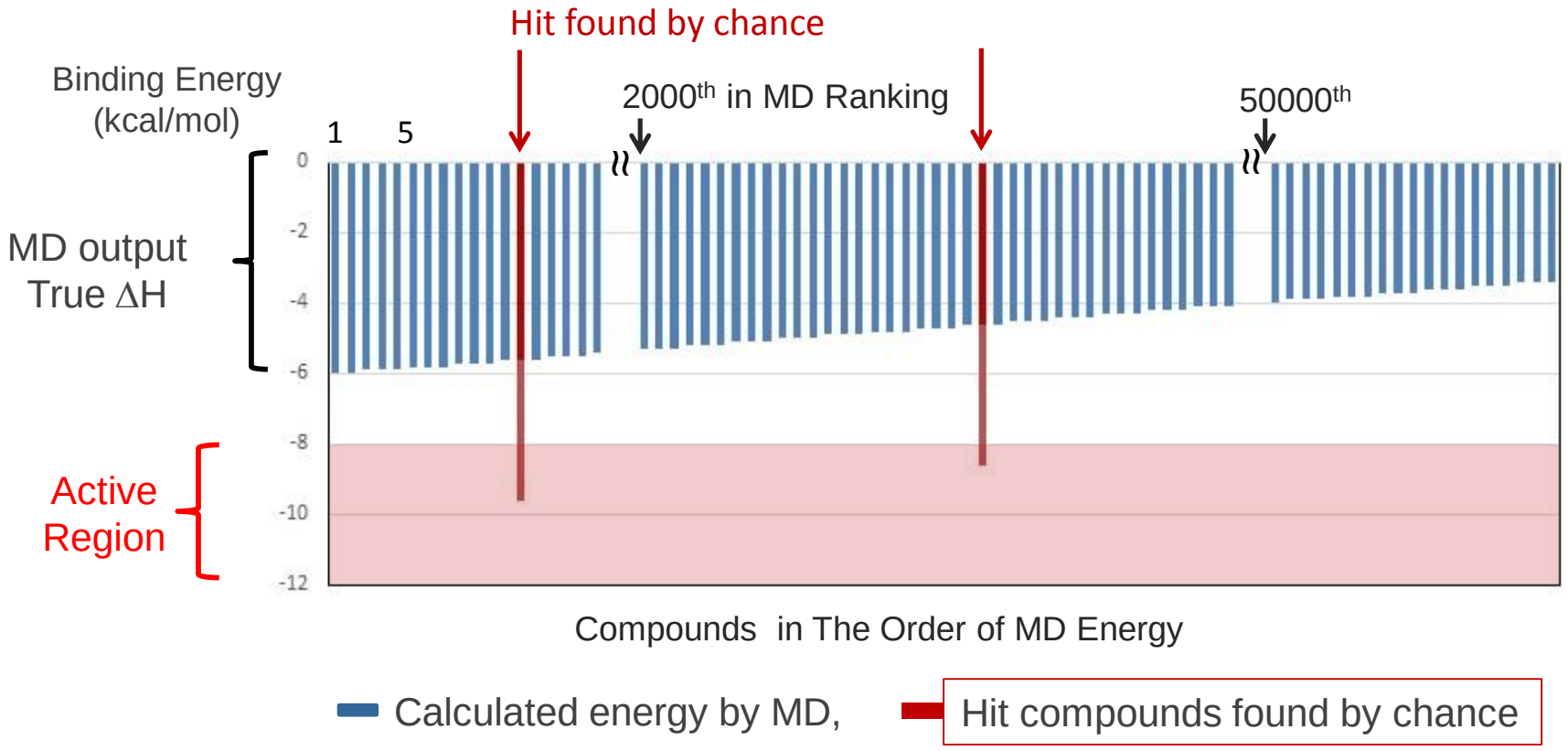
In MD-based docking methods, binding energy is ΔH and most of the proposed compounds do not show activity. This leads to very high false-positive rate.



Interprotein

Docking problem | After removal of assumed -TΔS

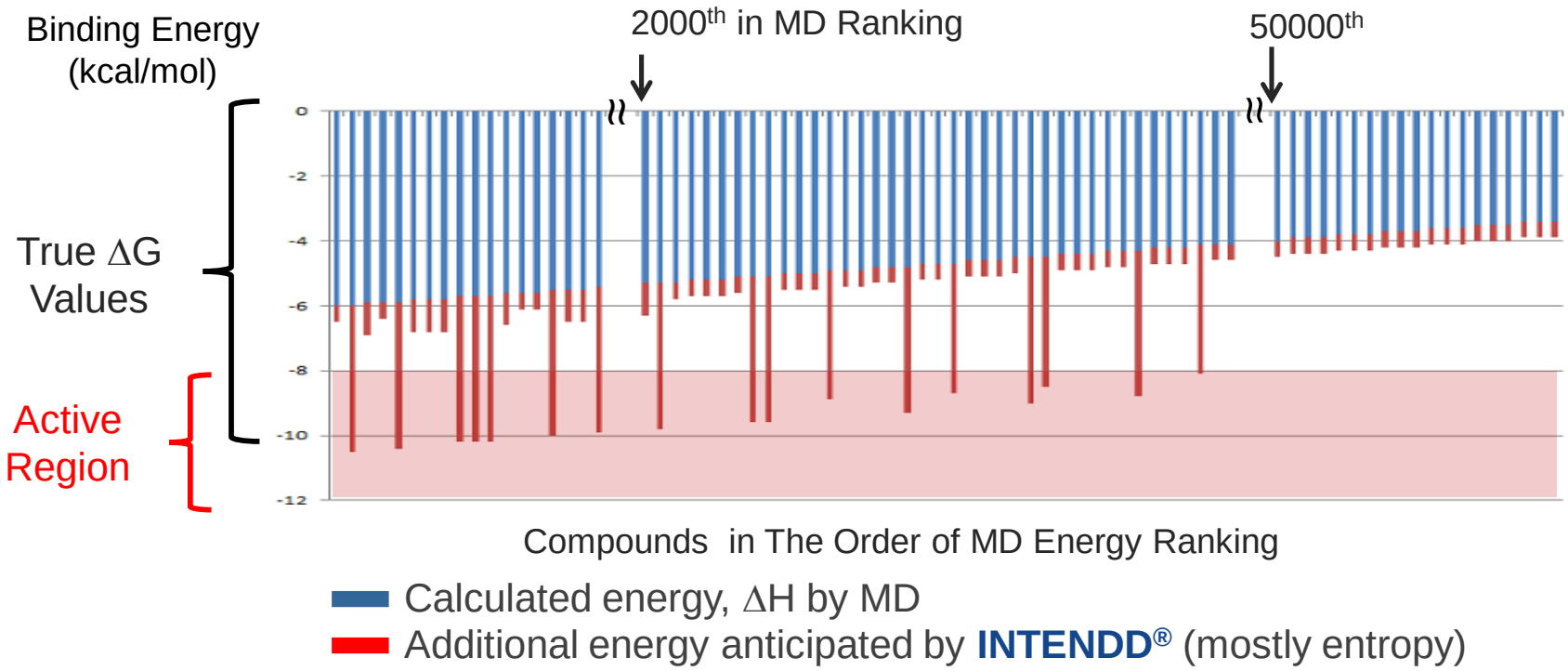
In MD-based docking methods, binding energy, ΔH is adjusted to observed ΔG. When entropy has "Push" effect and subtracted from ΔH, remaining ΔH is small and can not be a measure of activity.



Interprotein

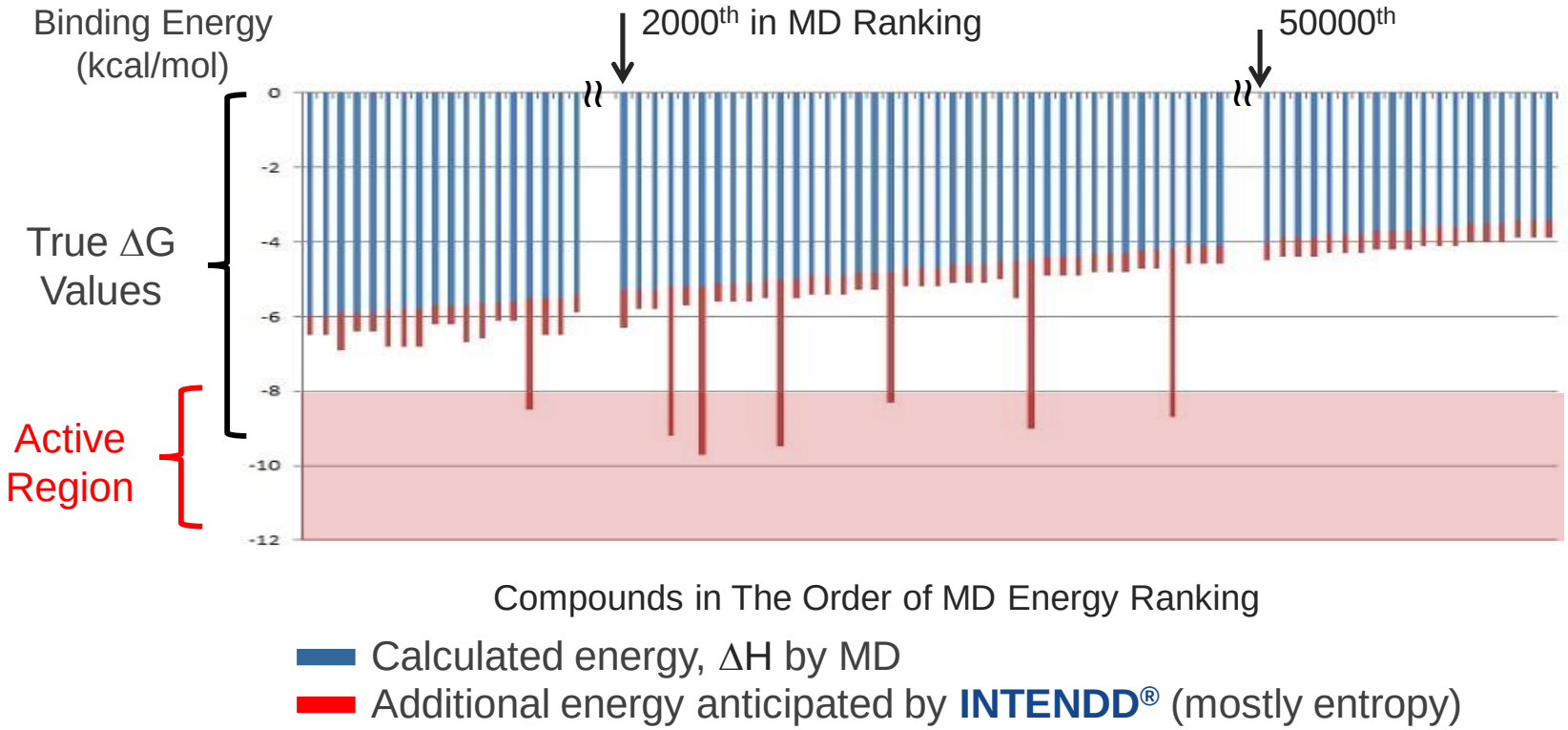
Kinase ATP sites are easy targets

ATP binding site of kinase is a flat pocket which accepts compounds with similar shape and size. Therefore, the contribution of entropy is similar among compounds. This is why the ΔH calculation can indicate hit compounds and population of hit in top group is high.



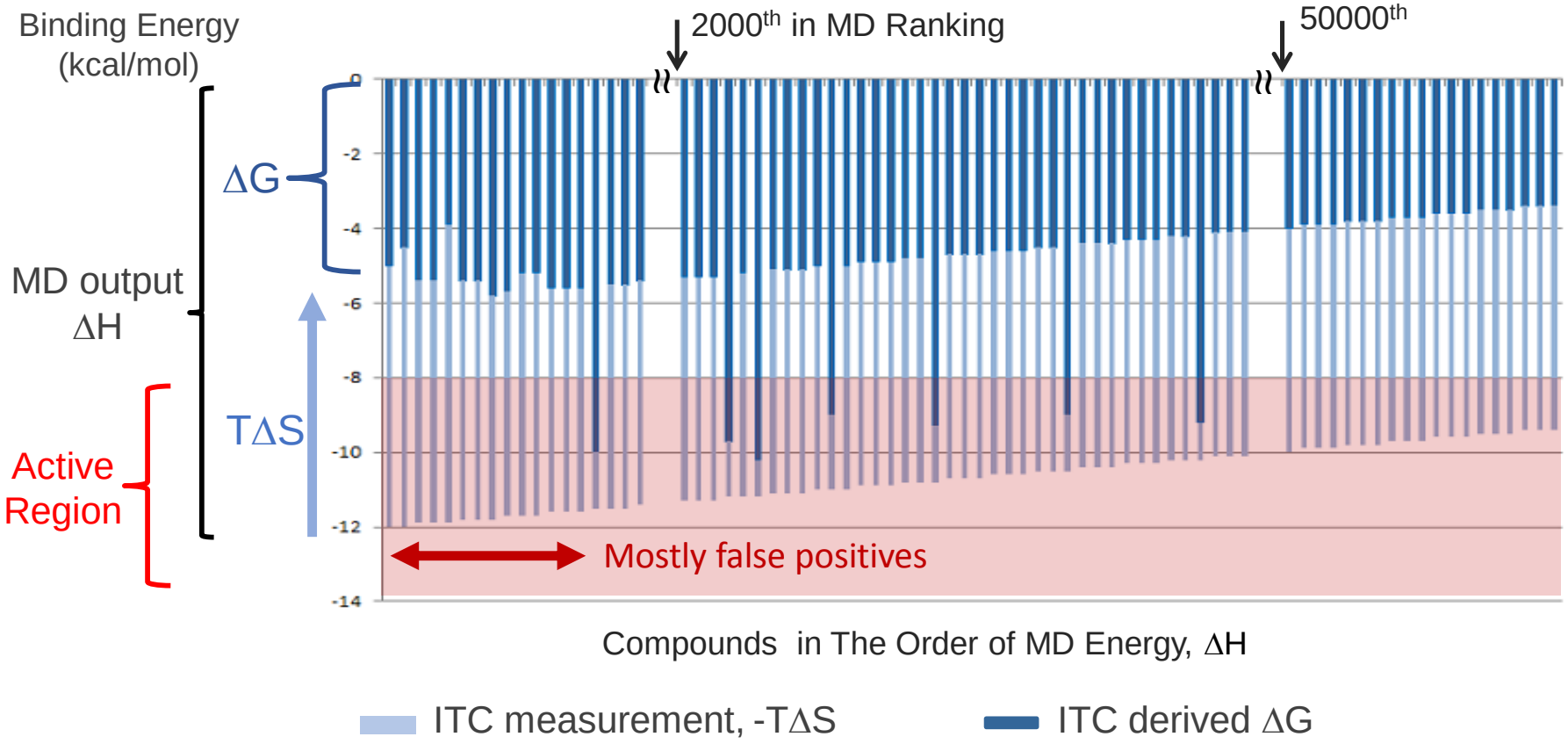
The real landscape of binding energy | ΔS "Push" case

- Activity Kd as $\Delta G = kT\ ln\ Kd$
- Binding Energy $\Delta G = \Delta H - T\Delta S$
- Docking calculation can estimate most of ΔH.
- But ΔS contribution is more significant than was expected.



Alternative Case | Case of ΔS "Pull"

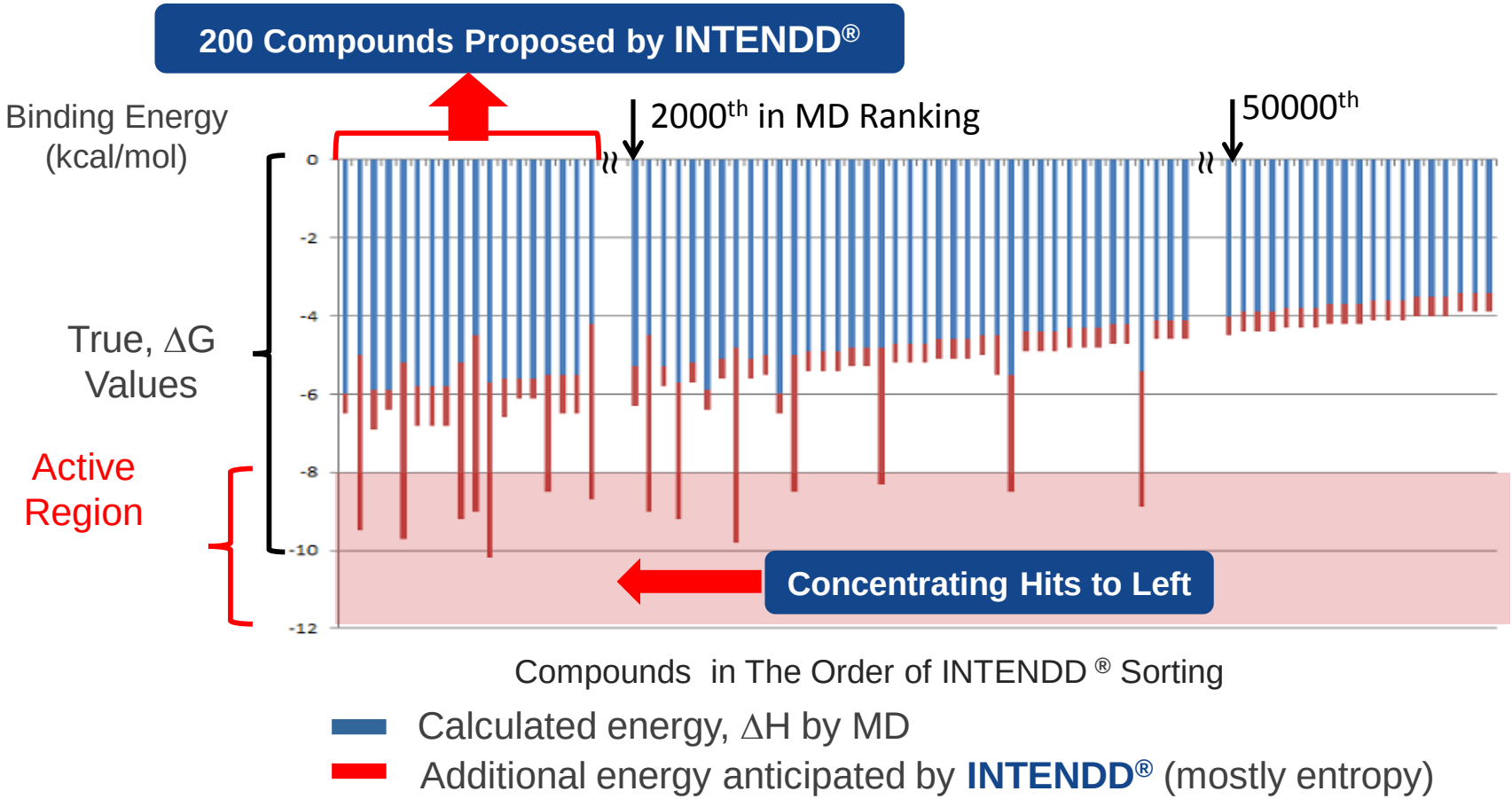
ITC measurements sometimes show negative entropy effect in drug-protein binding. Entropy effect is quite big and most compounds loose binding affinity because of decreased entropy.



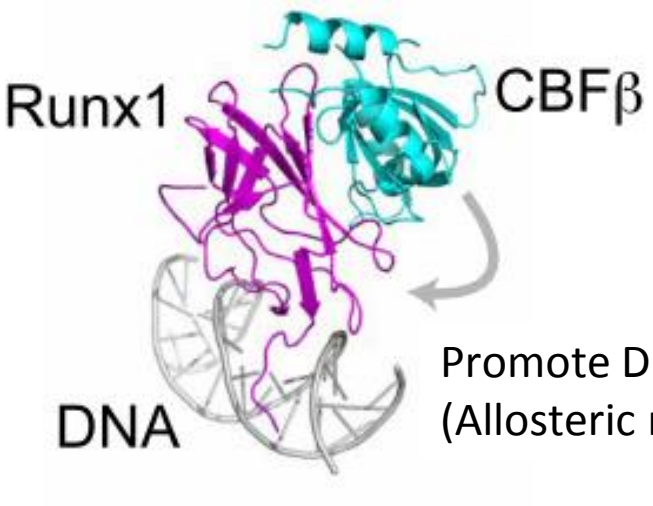
Interprotein

What our SBDD system, INTENDD® does.

INTENDD® takes entropy into account, bringing promising compounds with big entropy to higher ranks.



Strategy for Runx1 Inhibitor

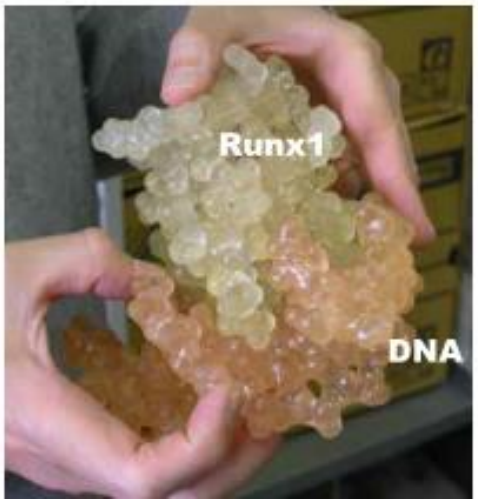


Promote DNA binding
(Allosteric modulation)

Runx1-CBFβ
Inhibitor

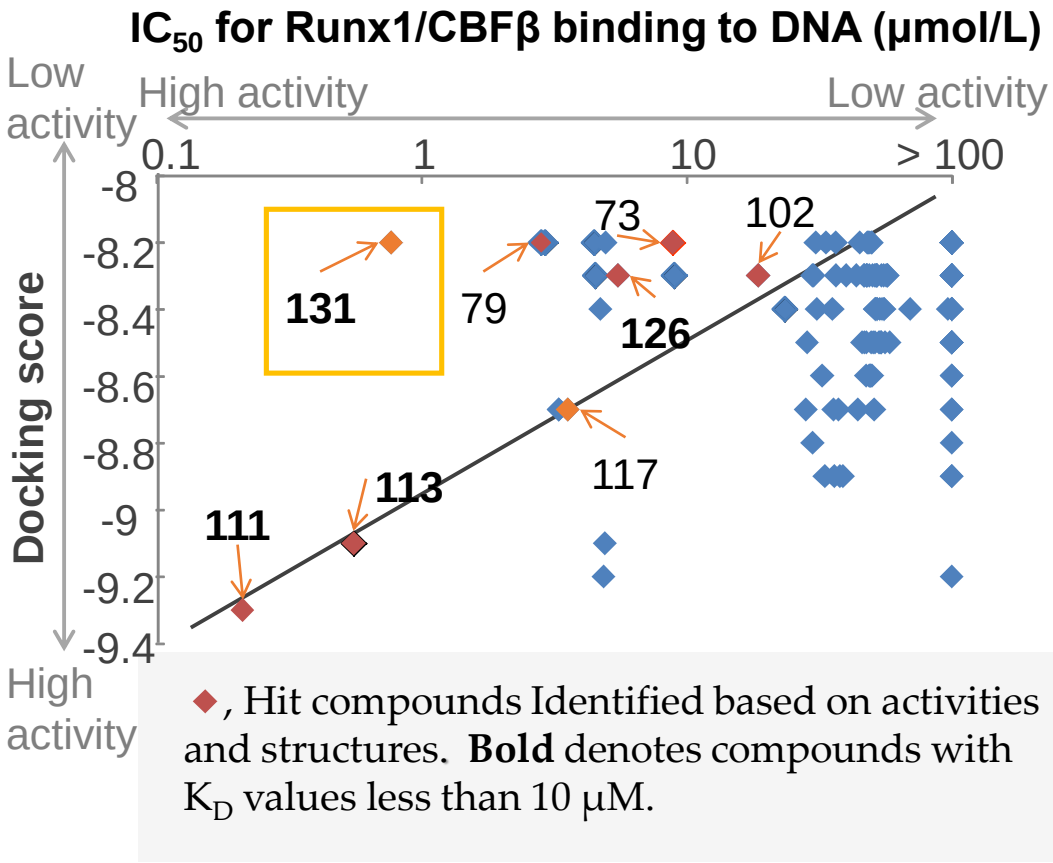
Runx1-DNA
Inhibitor

Drug discovery for Runx1 Inhibitor by INTENDD®



Identify Hot Spot	Use of "Real Space 3D Models" to identify a hot spot and to find the information on interaction between ligand and protein atoms
Ring Position	Search of ring positions at a hot spot
Ca. 50 skeletons	Formation of ligand's molecular skeletons (ca.50 structures)
Cmps Data Base	Search skeletons in compounds data bases and discover compounds that are expected to bind to the target protein
Collision Check	Collision check with protein atoms
Selection	Selection of promising compounds by clustering and filtering through binding structure-based mechanism
Output	Ca. 200 compounds to be provided for wet screening

Relationship between docking score and activity



◆, Hit compounds Identified based on activities and structures. **Bold** denotes compounds with K_D values less than 10 μM.

Comps.	kcal /mol	IC_{50} (μM)	K_D (μM)
73	-8.2	8.9	ND
79	-8.2	2.9	35
102	-8.3	19	ND
111	-9.3	0.21	5.2
113	-9.1	0.55	1.6
117	-8.7	3.6	ND
126	-8.3	5.5	0.065
131	-8.2	0.77	9.5

ND: Not determined.

There is no linear relationship between MD-based docking score and activity. For instance, the lower right field contains comps with high docking scores and low activities, while the upper left field contains comps with low docking scores and high activities. **INTENDD®** can detect highly active comps even among comps with low docking scores (**131**). **131** analogs showed a good SAR, resulting in the production of many highly active comps (IC50 < 1μM) in second screening.

Practical example | Runx1-CBFβ interaction Inhibitor

Primary hits (IC_{50} < 100 μM)

Number of hit compounds that inhibit Runx1/CBFβ-DNA binding over 50%

- 26 of 142* (hit rate: 18%) at the concentration of 100 μM
- 7 of 142 (hit rate: 5%) at the concentration of 10 μM
- 3 of 142 (hit rate: 2%) at the concentration of 1 μM

* Number of compounds tested following proposal by INTENDD®/SBSG®

1) Inhibition of Runx1/CBFβ-DNA binding by compounds was assessed by SPR.
2) Binding affinity of compounds to Runx1 was determined by MST.
3) ND: not determined.

Compound	73	79	102	111	113	117	126	131
IC_{50} for Runx1/CBFβ-DNA binding (μM)[1)]	8.9	2.9	19	0.21	0.55	3.6	5.5	0.77
Binding affinity to Runx1(K_D, μM)[2)]	ND[3)]	35	ND	5.2	1.6	ND	0.065	9.5

Number of tested compounds in secondary screening	1	0	2	2	6	10	17	35	Total 73
Number of secondary hit compounds	1	-	0	0	6	6	1	20	34
Hit rate in secondary screening (%)	100	-	0	0	100	60	5.9	57	46

Secondary hits (IC_{50} < 10 μM)