Spresso:
An ultrafast compound pre-screening method based on compound decomposition

Keisuke Yanagisawa[1,2]
Shogo D. Suzuki[2,3]
Takashi Ishida[1,2,3,4]
Shunta Komine[2,3]
Masahito Ohue[1,3,4]
Yutaka Akiyama[1,2,3,4]

1. Dept. CS, School of Computing, Tokyo Tech.
2. Education Academy of Computational Life Sciences (ACLS), Tokyo Tech.
3. Dept. CS, Graduate School of Information Science and Engineering, Tokyo Tech.
4. Advanced Computational Drug Discovery Unit (ACDD), Tokyo Tech.

東京工業大学
Tokyo Institute of Technology

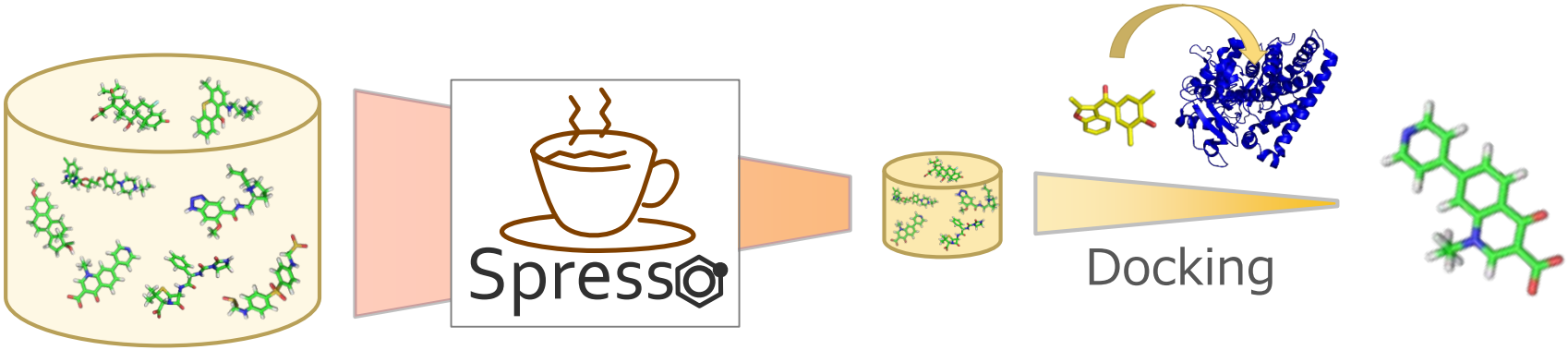
Education Academy of Computational Life Sciences

* The software was renamed from ESPRESSO to Spresso (same pronunciation) in October 2016.

Spresso

Speedy **PRE**-**S**creening method with **S**egmented c**O**mpounds

(http://www.bi.cs.titech.ac.jp/spresso/)



> 10^7 entries

~ 10^5 entries

~ 10^3 entries

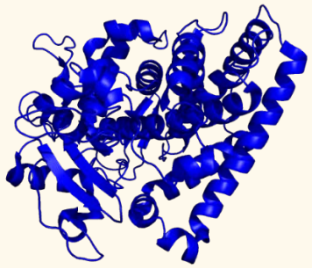
Docking-based virtual screening

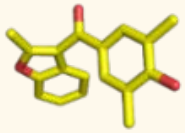
Virtual screening

Compound DB

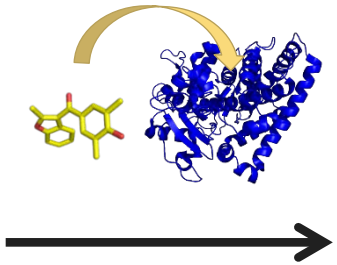
Drug candidates

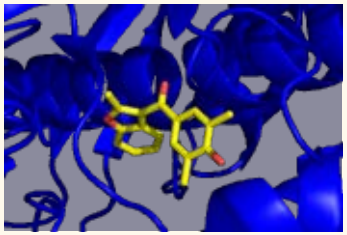
Docking calculation





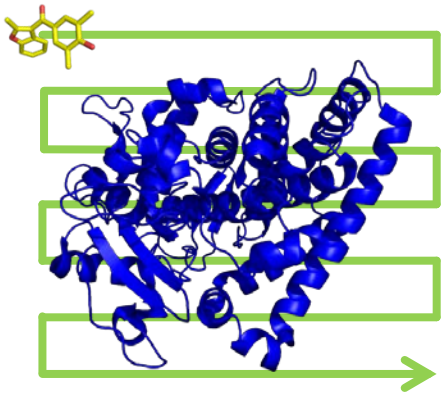
A protein & a compound



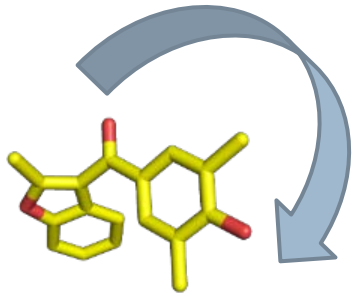


A conformation

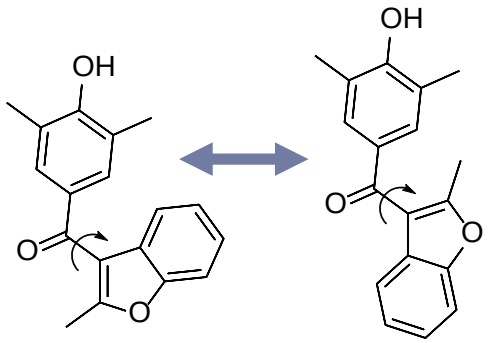
Conformation search



Translation
3 dimensions



Rotation
3 dimensions

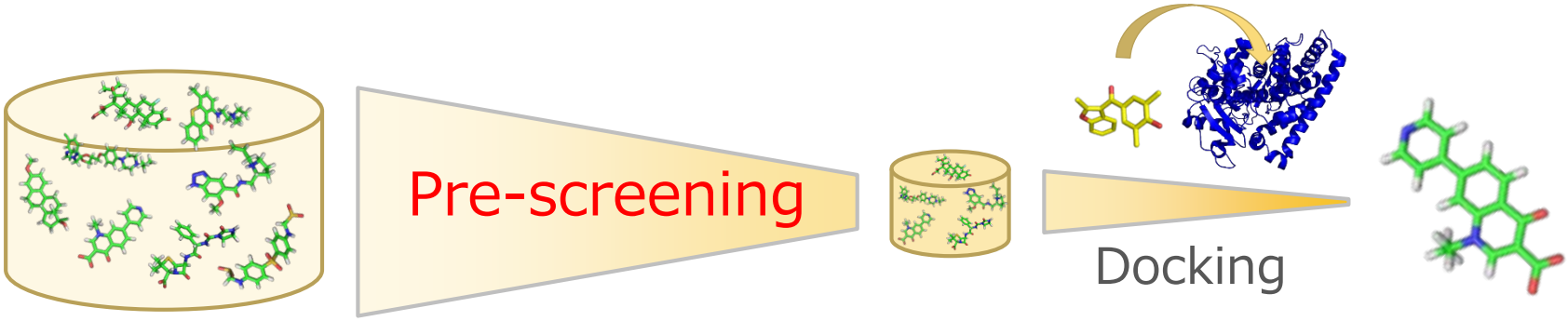


Internal rotation
N dimensions

Problem: **Computationally expensive**

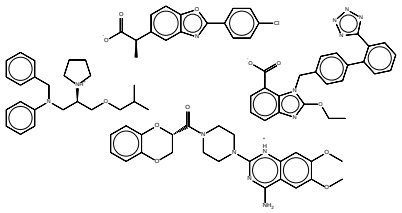
Compound pre-screening

Decreasing calculation with pre-screening

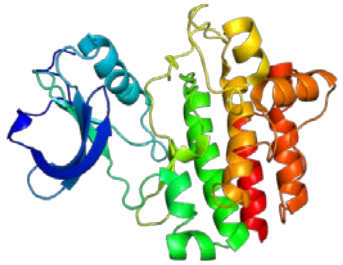


Ex: 10,000,000 compounds → (x 1/100) → 100,000 compounds → 1,000 compounds

Existing pre-screening methods



Pre-screening



Ligand-based

ML[1], Shape matching[2] etc.

☹ Requiring known compounds
resulting in difficulty finding
novel drug candidates

☺ Light weight calculation

Structure-based

Rough docking (Glide HTVS)[3,4]

☺ No bias toward known actives

☹ **Insufficient speed**
(1 sec / compound =
1 CPU year for ZINC DB)

1) H. Geppert, *et al*. *J Chem Inf Model*, 2010.
2) A. Vuorinen, *et al*. *J Med Chem*, 2014.
3) S. B. Mirza, *et al. J Mol Graph Model*, 2016.
4) A. Grover, *et al*. *Biochim Biophys Acta*, 2012.

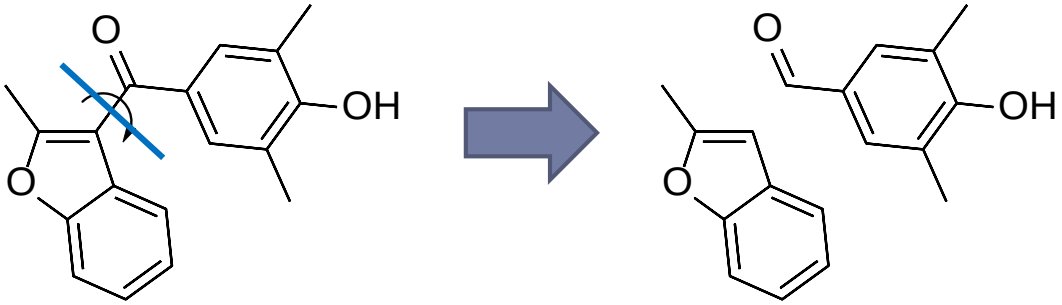
Our approach

- **Structure-based**
- **Ultrafast** compared to existing method

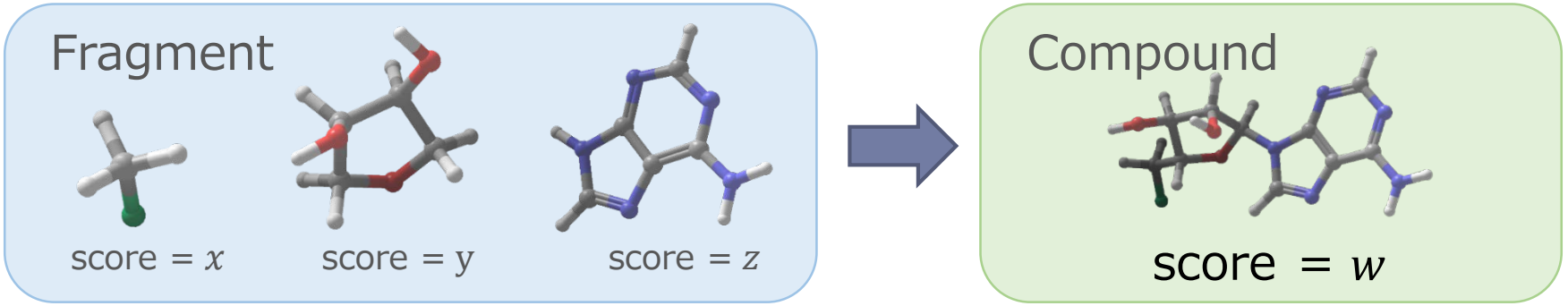
	Ligand-based	Structure-based	Spresso
wo/ known compound	☹	☺	☺
pre-screening speed	☺	☹	☺

Ideas for acceleration

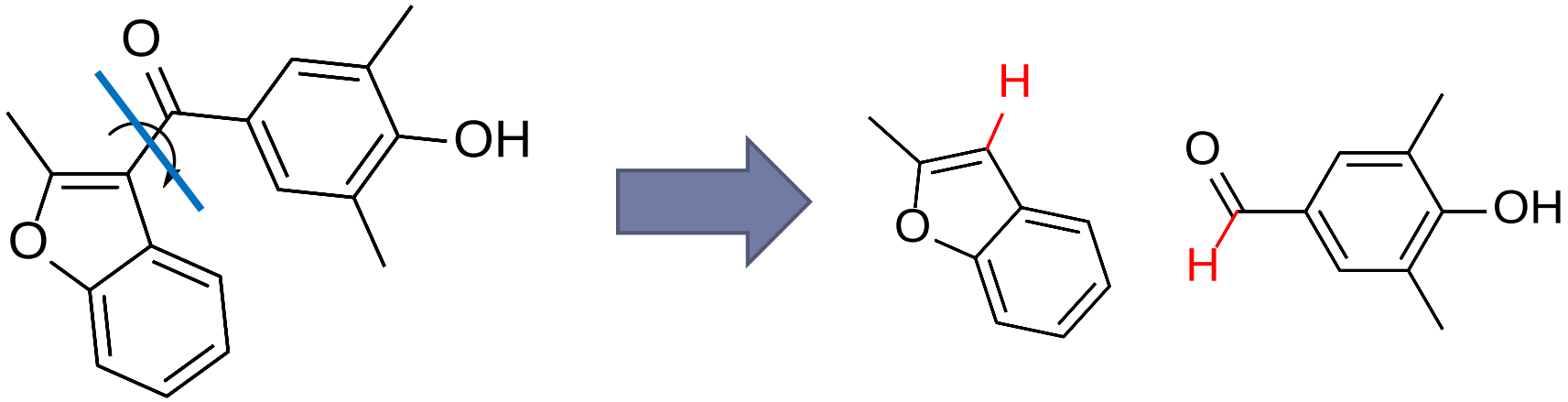
Idea I. Compound decomposition



Idea II. Rough compound evaluation



Idea I. Compound decomposition



Compound

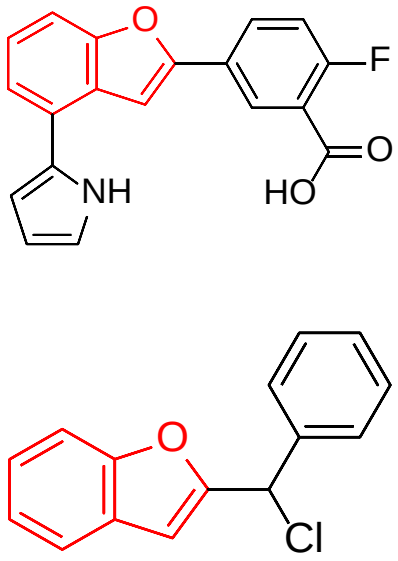
Fragments

Creating fragments without any rotatable bond[1)]

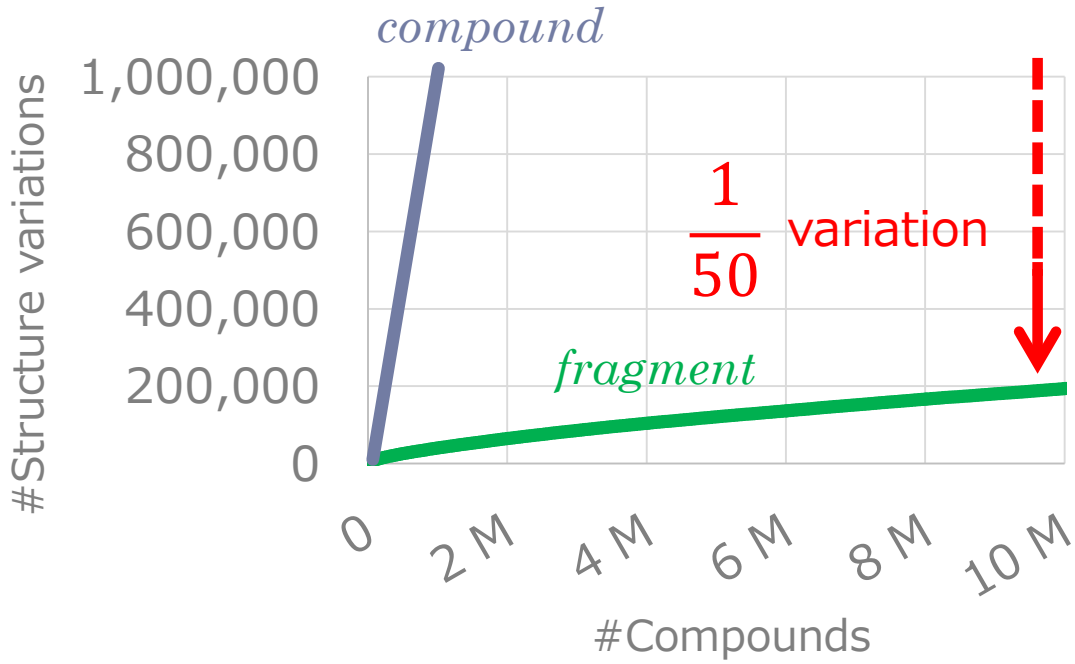
1) S. Komine *et al.*, *IPSJ SIG Technical Report*, 2015-BIO-42, 2015.

Another benefit of decomposition

Sharing of fragment docking results for duplication



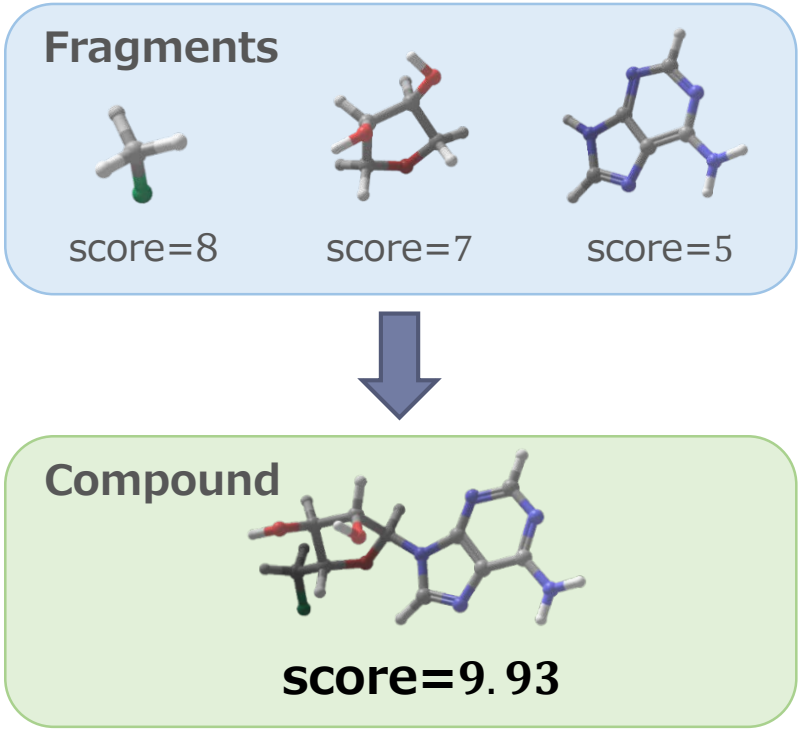
Example of
duplicated fragments



Decomposition result
compounds: ZINC “drugs now” subset (10,639,555 entries)

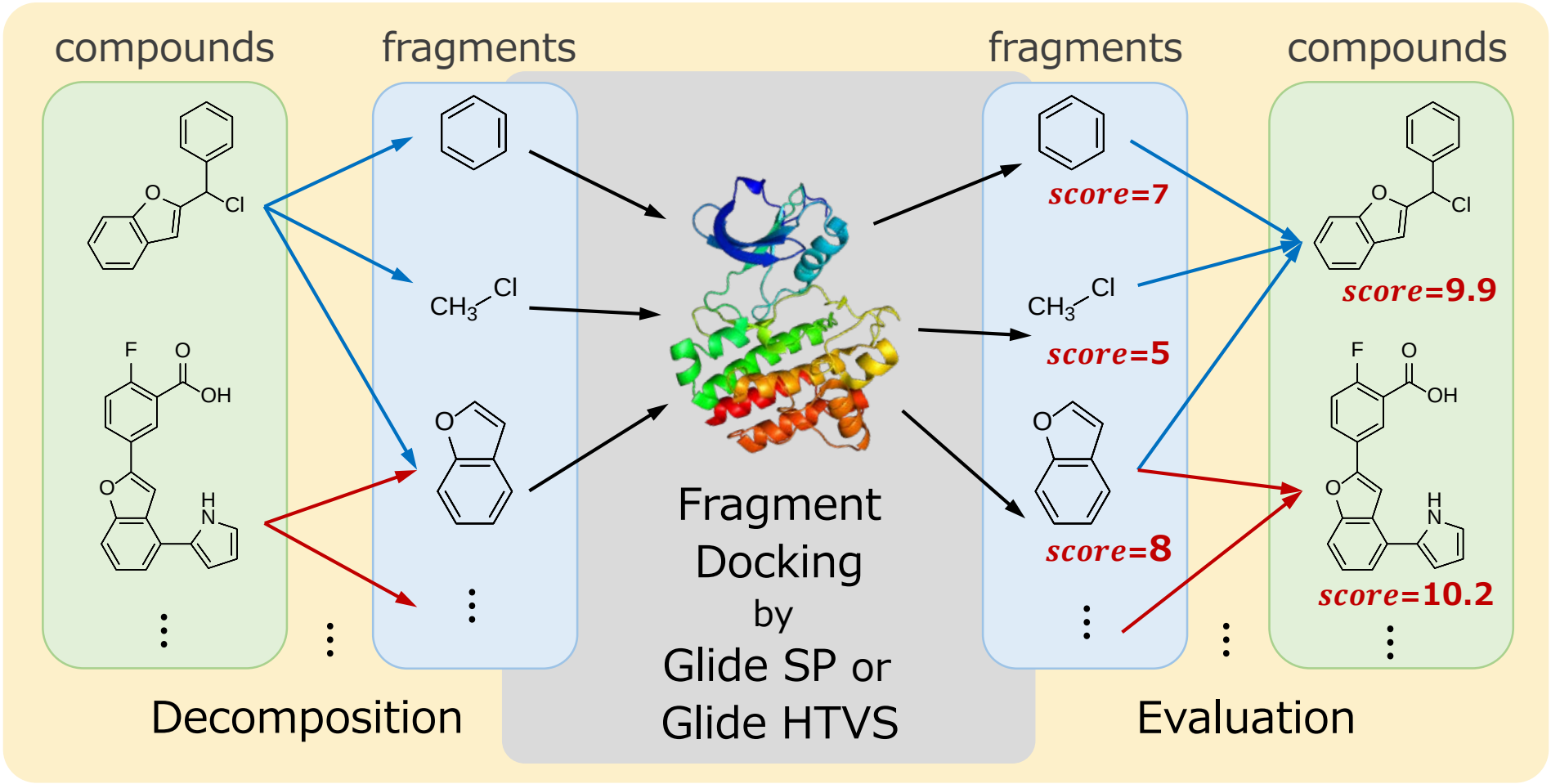
Idea II. Rough compound evaluation

Compound evaluation without re-construction



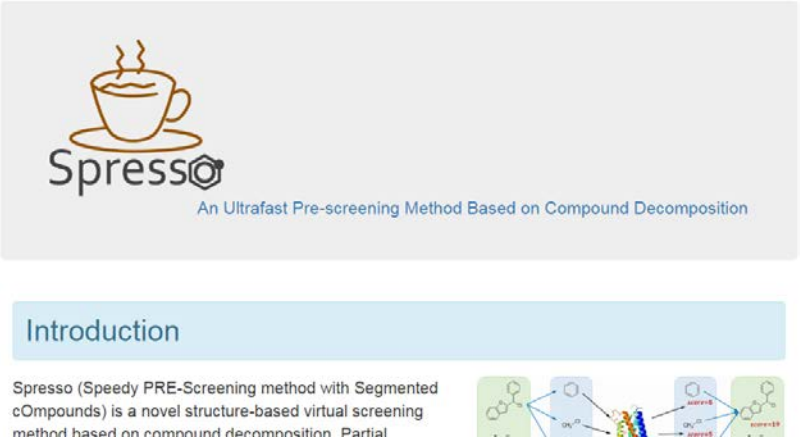
Generalized Sum-3 (GS_3)
of fragment scores is adopted

$$\mathrm{GS}_3 = \sqrt[3]{\sum_f (\mathrm{score}_f)^3}$$



Spresso is **open-sourced** under GPLv3 license

(http://www.bi.cs.titech.ac.jp/spresso)



Acknowledgements

東京工業大学
Tokyo Institute of Technology

東京工業大学 情報生命博士教育院
Education Academy of Computational Life Sciences

国立研究開発法人
科学技術振興機構
Japan Science and Technology Agency

科研費
KAKENHI