

# Toward Crowdsourcing Evaluation of Synthetic Accessibility

Yukino Baba<sup>1</sup>, Tetsu Isomura<sup>2</sup>,  
Hisashi Kashima<sup>1</sup>, Hirofumi Nakano<sup>1</sup>

1: Univ. of Tokyo  
2: The KAITEKI Institute, Inc.

## OVERVIEW

### We perform a feasibility study of crowdsourcing evaluation of synthetic accessibility

- Screening of compounds according to their synthetic accessibility has key importance in drug discovery process.
- We anticipate that *crowdsourcing*, a platform for outsourcing microtasks to a large group of people, would accomplish the rapid assessment of synthetic accessibility.

## OUR GOAL

### Building a drug discovery process where crowdsourcing evaluation of synthesis accessibility is incorporated

## EXPERIMENTS

### Participants

4 experts, 4 semi-experts, 100 non-experts

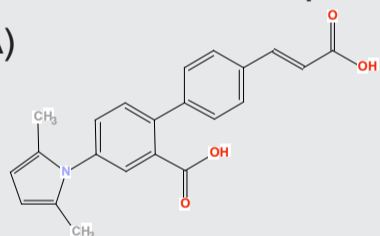
### Compound sets

Set A: 10 already existing compounds

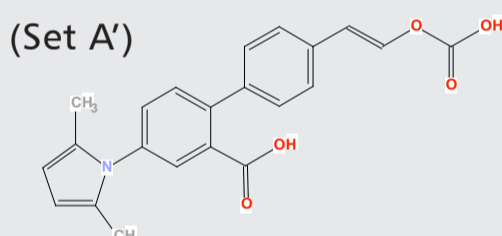
Set A': 10 hard-to-be-synthesized compounds (artificially generated from Set A)

Set B: 10 novel compounds

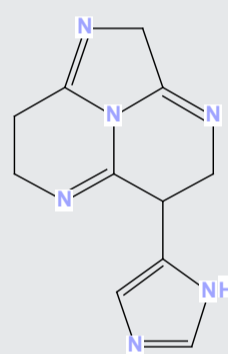
(Set A)



(Set A')



### Instructions



Suppose that you are asked to synthesize this compound in three months. Please rate the ease of synthesis on a five-point scale.

- 1: Difficult
- 2: Slightly difficult
- 3: Neutral
- 4: Slightly easy
- 5: Easy

## RESULTS

(1) The experts accurately evaluated the synthetic accessibility even if they only spent 15-30 seconds per compound on average.

(2) The non-experts did not show a good performance on the synthetic accessibility assessment. Not so much as the experts, but the semi-experts evaluated the synthetic accessibility more accurately than the non-experts

Ratio of correct answers

Expert A	Expert B	Expert C	Expert D	Semi-expert E	Semi-expert F	Semi-expert G	Semi-expert H
1.00	0.95	0.80	0.90	0.45	0.65	0.55	0.75
Average: 0.91				Average: 0.60			
Average of non-experts: 0.52							

\* Correct answer:

When a participant gives lower rate to a compound from the hard-to-be-synthesized set (Set A') than the original one in Set A, the answer is considered as *correct*.

Level of agreement between experts and others

	Set A	Set A'	Set B	Avg.
Avg. distance of responses between each pair of experts	0.954	0.619	0.789	0.787
Avg. distance of responses between each pair of expert and semi-expert	1.440	1.071	0.951	1.154
Avg. distance of responses between each pair of expert and non-expert	1.515	1.092	1.331	1.313

Average response time per compound (sec)

Expert A	Expert B	Expert C	Expert D	Semi-expert E	Semi-expert F	Semi-expert G	Semi-expert H
17.3	14.5	29.7	28.3	22.3	52.7	151.6	15.4
Average: 22.5				Average: 60.5			