

ROUTE SCOUTING

(aka *RETROSYNTHETIC ANALYSIS*)

<https://en.wikipedia.org> › wiki ▾ Översätt den här sidan

Retrosynthetic analysis - Wikipedia

Retrosynthetic analysis is a **technique for solving problems in the planning of organic syntheses**. This is achieved by transforming a target molecule into ...

[Definitions](#) · [Example](#) · [Strategies](#)

<https://en.wikipedia.org> › wiki ▾ Översätt den här sidan

Retrosynthetic analysis - Wikipedia

Retrosynthetic analysis is a technique for solving problems in the planning of organic syntheses. This is achieved by transforming a target molecule into ...

[Definitions](#) · [Example](#) · [Strategies](#)

<https://chemistnotes.com> › organic ▾ Översätt den här sidan

Retrosynthetic analysis:Definition, example - Chemistry Notes

7 juli 2022 — Retrosynthetic analysis is an imaginative procedure in which the target molecule(TM) is broken down or disassembled into less complicated, ...

ROUTE SCOUTING

(aka *RETROSYNTHETIC ANALYSIS*)

ROUTE SCOUTING REVOLUTION

revolution

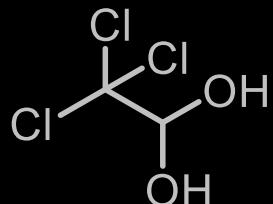
revolution

“...a sudden and great change...”

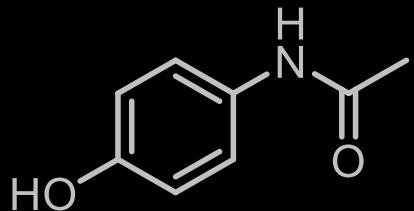
<https://dictionary.cambridge.org/dictionary/english/revolution>



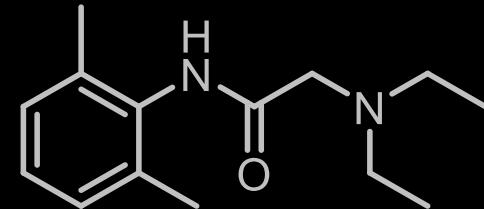
drug structures: evolution



CHLORAL HYDRATE
(1832)



PARACETAMOL
(1877)



LIDOCAINE
(1943)



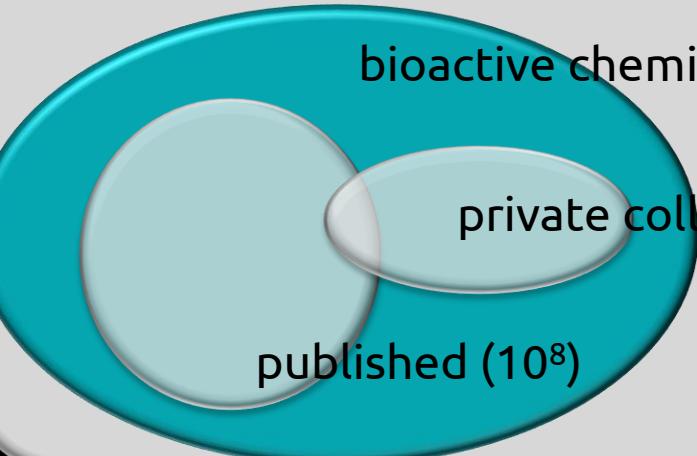
LINAGLIPTIN
(2006)



UMECLIDINIUM
(2014)

drug structures: evolution

chemical space (10^{180})



human genome ca 30.000

druggable genome ca 3000

drug targets ca 600-1500

disease-modifying genome ca 3000

the easy ones already exist

retrosynthetic analysis tools



Robert B. Woodward
(1917-1979)
(Nobel prize 1965)



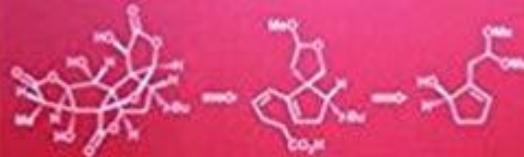
Alexander R. Todd
(1907-1997)
(Nobel prize 1957)

retrosynthetic analysis is art



Elias J. Corey
(b. 1928)
(Nobel Prize 1990)

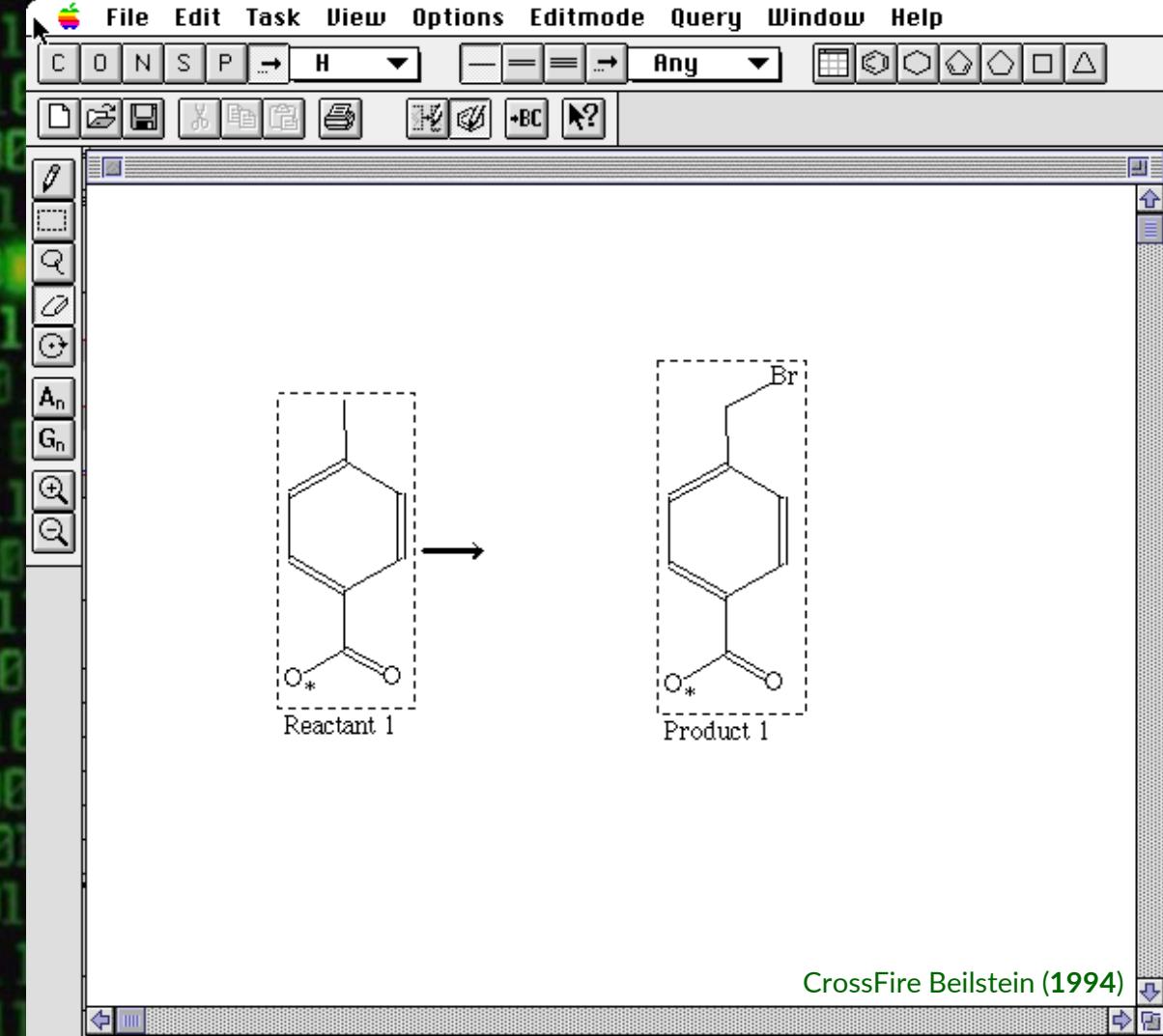
The LOGIC of CHEMICAL SYNTHESIS



*E.J. Corey &
Xue-Min Cheng*

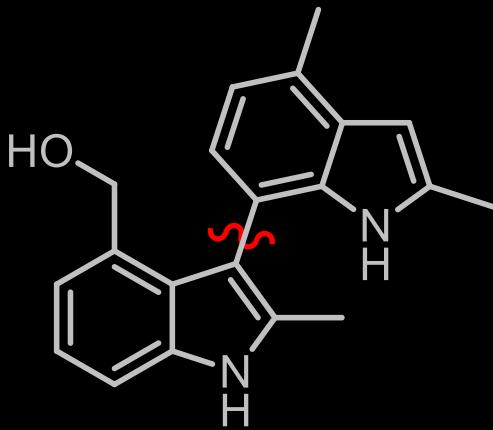
ROUTE SCOUTING

FIRST REVOLUTION



how is it done
today?

target molecule



(2,2',4'-trimethyl-1H,1'H-[3,7'-biindol]-4-yl)methanol

ROUTE SCOUTING TOOLS

i&os

Saved
and Alerts

History



Searching for...

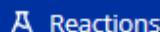
Reactions

Search by Substance Name, CAS RN, Patent Number, PubMed ID, AN, CAN, and/or DOI. [Learn More](#)

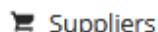
Enter a query...



Substances



References



Biosequences

Retrosynthesis

CAS Draw

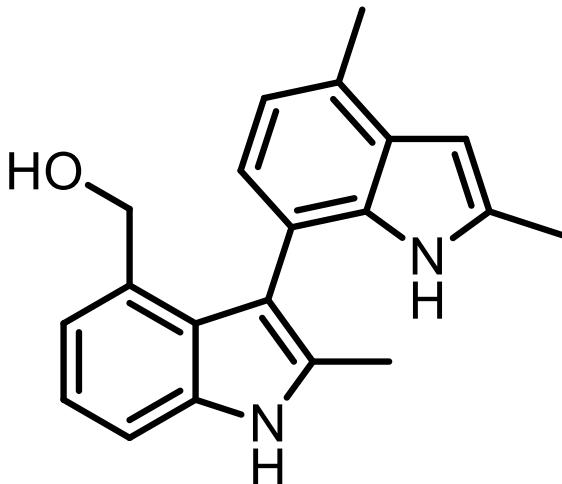
X



Enter a CAS Registry Number, SMILES, or InChI...



Draw or change atoms or bonds.



Molecular Formula:

C



Zoom: 100%

OK

Cancel

Sea

oo A

C Et

X R

A R

+ -

[]₁₋₄

S

O_n

A B

1-1

E

R

→

Recent Search His

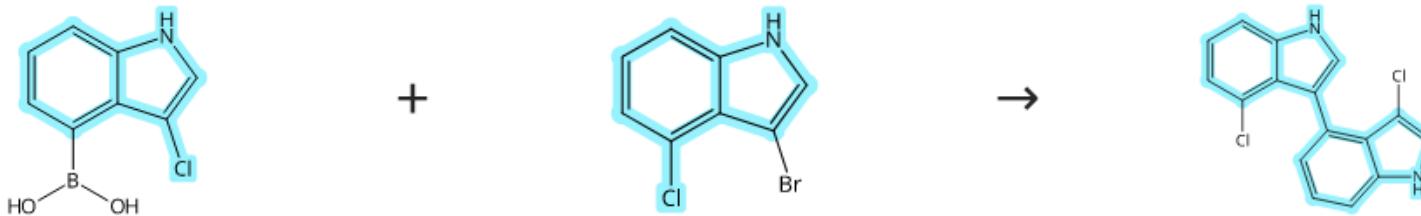
September 29

6:52 AM

C	H
O	S
N	P
Cl	Si
\	/
\	=E
pentagon	pentagon
hexagon	hexagon
heptagon	heptagon

Scheme 1 (1 Reaction)

Steps: 1 Yield: 55%

 Suppliers (34)

Reaction Summary

Steps: 1 Yield: 55%

1.1 Reagents: Potassium carbonate

Catalysts: Tetrakis(triphenylphosphine)palladium

Solvents: Tetrahydrofuran, Water; 5 h, 70 °C

Preparation of fused indole derivatives as organic electronic device materials

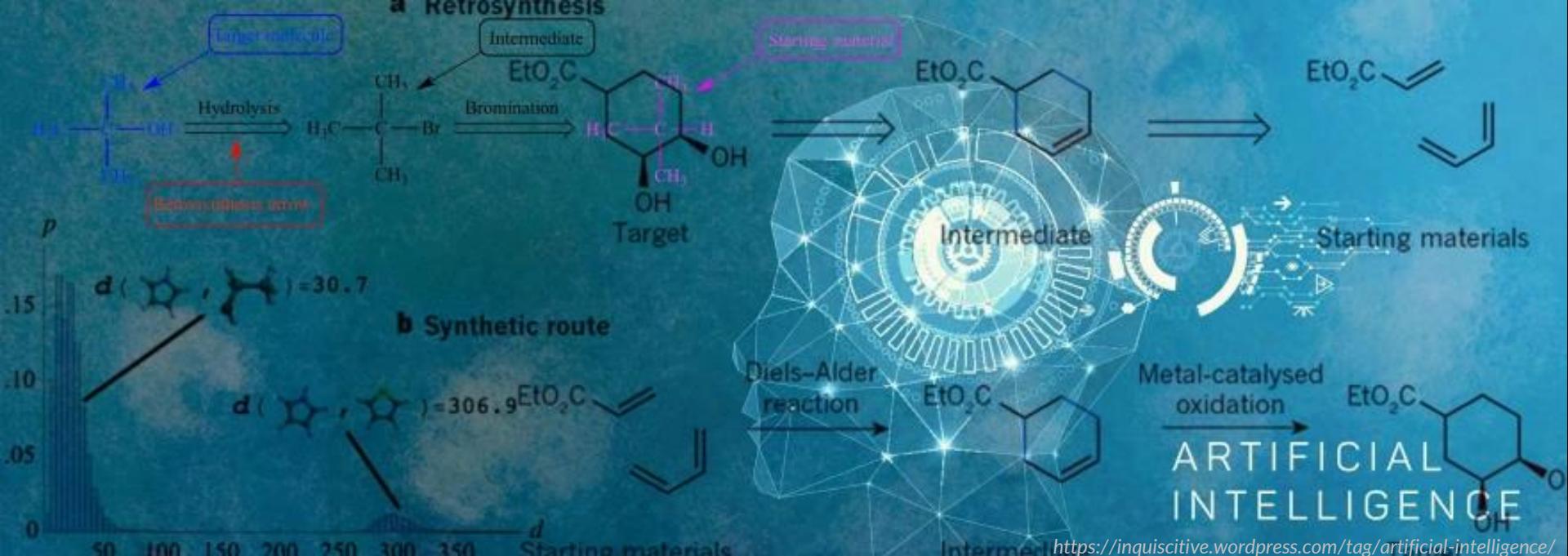
By: Choi, Don Su

Korea, Republic of, KR2175379 B1 2020-11-09

[View Reaction Detail](#)[PATENTPAK ▾](#)[Full Text ▾](#)

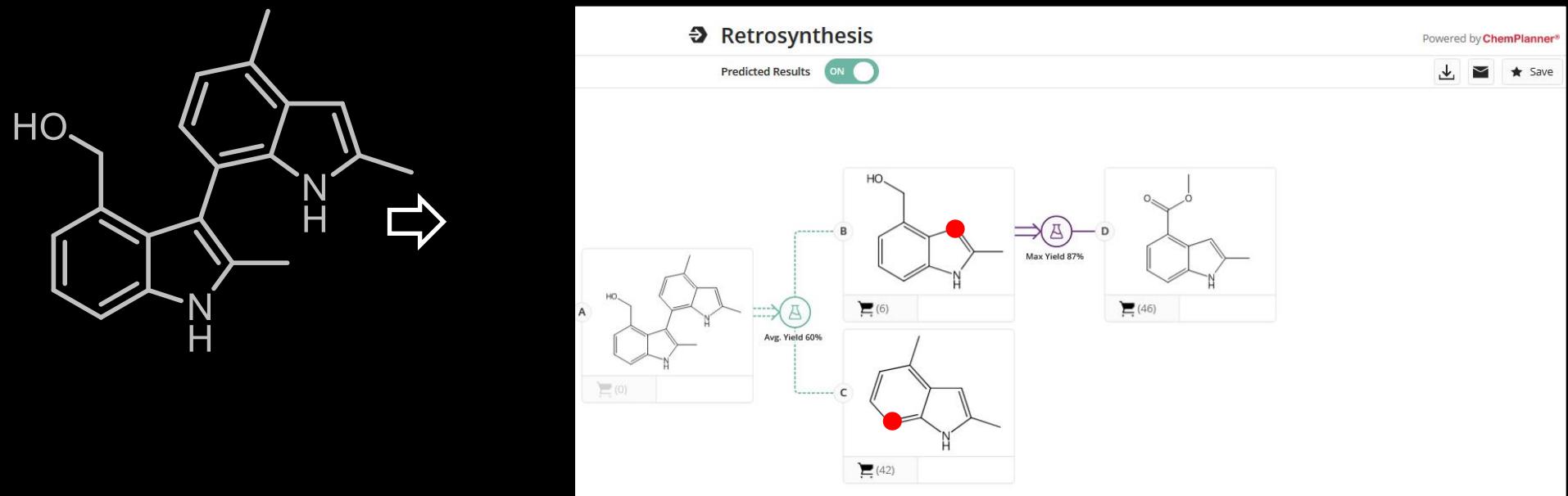
ROUTE SCOUTING REVOLUTION

“a sudden and great change”



computer aided synthesis design (CASD)

- LHASA
- SYNCHEM
- COMPASS
- WODCA
- SST
- KOSP
- HORACE
- AIZynthfinder
- SECS
- FLAMINGOES
- EROS
- SYNGEN
- CHIRON
- ARChem / ChemPlanner
- ICSYNTH
- ASKCOS



ROUTE SCOUTING TOOLS

i&os

The image shows a retrosynthetic analysis of a complex organic molecule. On the left, a large molecule is shown with an arrow pointing to a retrosynthetic interface. The interface has a header 'Retrosynthesis' with a 'Predicted Results' toggle set to 'ON'. It includes download, email, and save icons. The main area is titled 'AlZynthfinder' (test version) and displays 'Compounds to Procure' (Zinc entries): a substituted indole, a substituted indole with a carbonyl, and a cyclic amide. Below is a 'Steps' section showing a synthesis route. The route starts with two building blocks: a substituted indole with a carbonyl (highlighted in green) and a cyclic amide with a bromine atom (highlighted in orange). These combine to form a intermediate with a bromine and a hydroxyl group (highlighted in orange). This intermediate then reacts with another substituted indole (highlighted in green) and a substituted indole with a hydroxymethyl group (highlighted in orange) to yield the final product.

ROUTE SCOUTING TOOLS

i&os

Retrosynthesis

Predicted Results

AIZynthfinder

(test version; limited “policies” and resource databases (e.g.: s/ms only ZincDB)). Time taken: < 5 mins

Compounds to Procure

ASKCOS

Open source software tools for organic synthesis

Site by MIT offering some of their published analysis approaches (based on Reaxys)

Interactive Path Planning:

Currently selected:

Precursor:

Group smile:

Rank	Score
1	0.571-0.608
2	0.247-0.264
3	0.124-0.134
4	0.037-0.047
5	0.007-0.017

Nodes can be further Expanded (not shown)

Smiles: Cc1[nH]c2ccccc2c(O)c2c1Br

Template score: 0.002

Plausible: 0.07

Num. Examples: 36642

Supposing template:

- Se1f14b6e0348032850950911 (131 examples)
- Se1f14b6e0348032850950951 (1279 examples)
- Se1f14b6e0348032850950954 (15714 examples)
- Se1f14b6e0348032850950954 (13471 examples)
- Se1f14b6e0348032850950954 (13471 examples)
- Se1f14b6e03480328509509513 (1839 examples)
- Se1f14b6e0348032850950974 (168 examples)

Currently selected: **Path #5**

Smiles: Cc1[nH]c2ccccc2c(O)c2c1Br

Evaluate reaction in new tab

Reaction details:

Reactionscore: -104830.716

Template score: 0.002

Plausible: 0.07

Num. Examples: 36642

Supposing template:

- Se1f14b6e0348032850950911 (131 examples)
- Se1f14b6e0348032850950951 (1279 examples)
- Se1f14b6e0348032850950954 (15714 examples)
- Se1f14b6e0348032850950954 (13471 examples)
- Se1f14b6e0348032850950954 (13471 examples)
- Se1f14b6e03480328509509513 (1839 examples)
- Se1f14b6e0348032850950974 (168 examples)

Recommendation 1

Reagents:

Catalysts:

Solvent(s):

Temperature: 99 Celsius

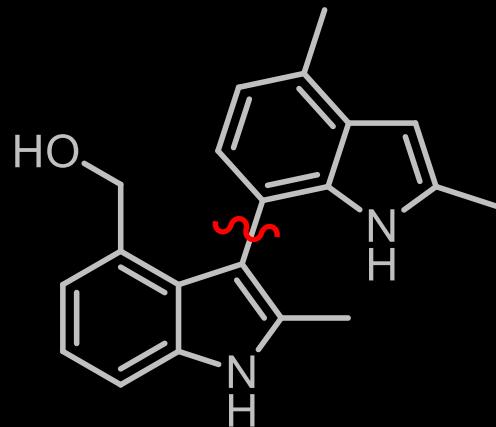
Recommendation 2

Synthetic Experts

SciFinder

ChemPlanner

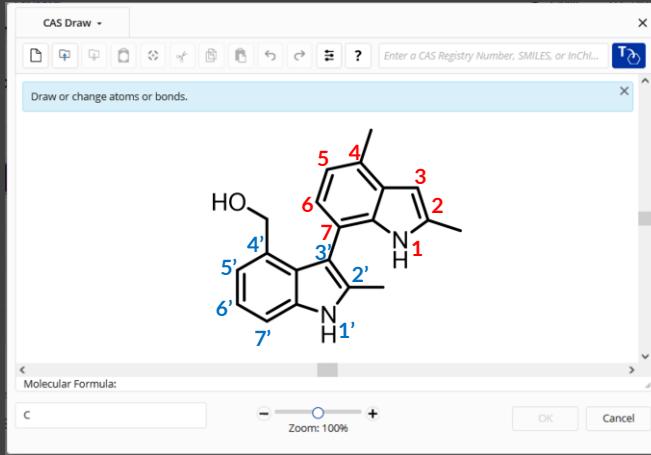
AI^Zynthfinder



ASKCOS

ROUTE SCOUTING TOOLS

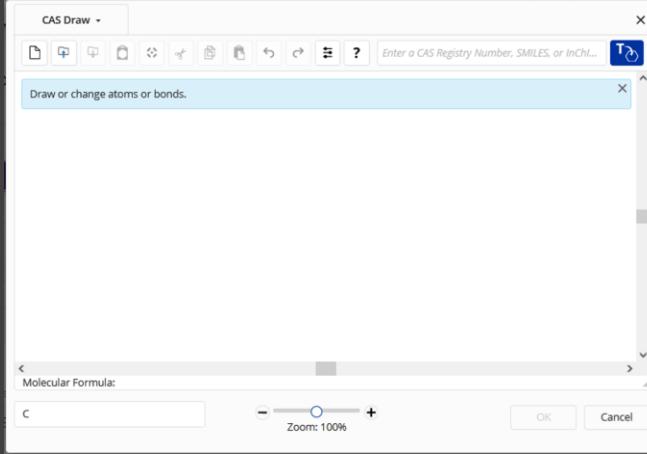
i&os



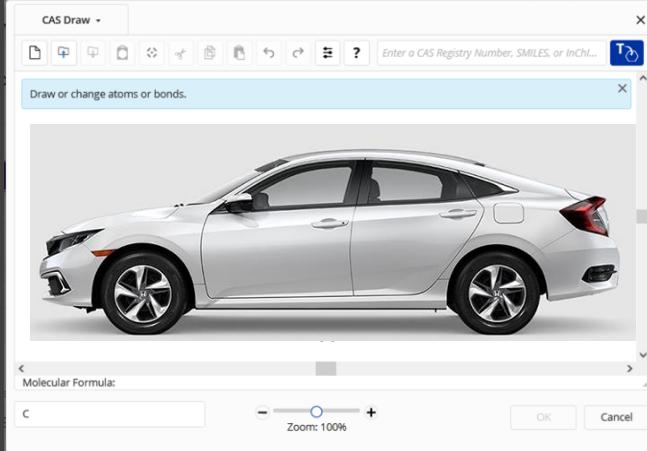
similarity based



*there's something about
similarity*



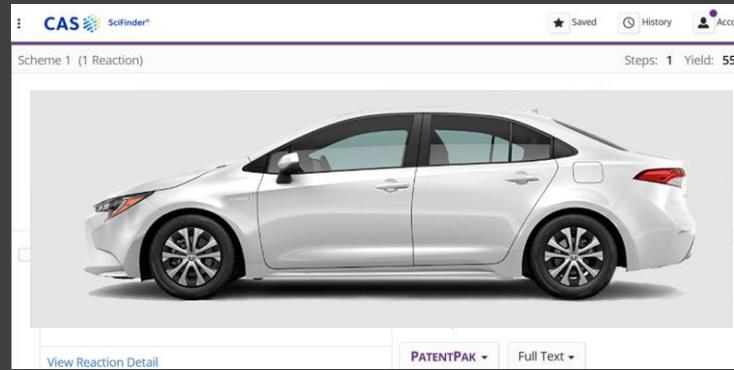
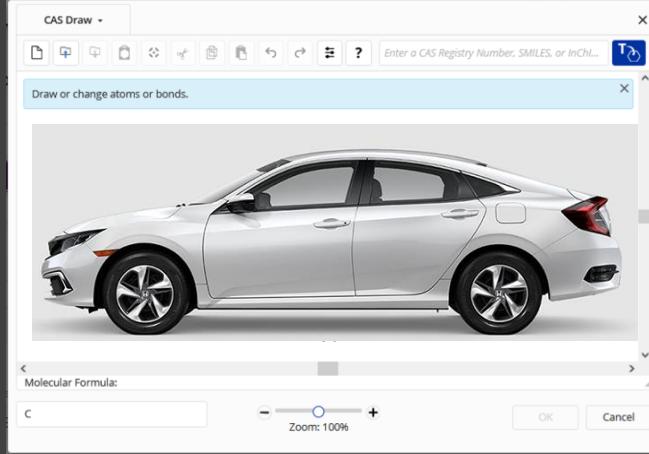
*there's something about
similarity*



*there's something about
similarity*

Similarity based on:

- SAME SHAPE (IMAGE)
- SAME FUNCTIONALITY



*there's something about
similarity*

Similarity "only" based on:

- SAME SHAPE (IMAGE)



*they look the same, but one of
them cannot sing*

SIMILARITY IS NOT ALL

(in organic synthesis)

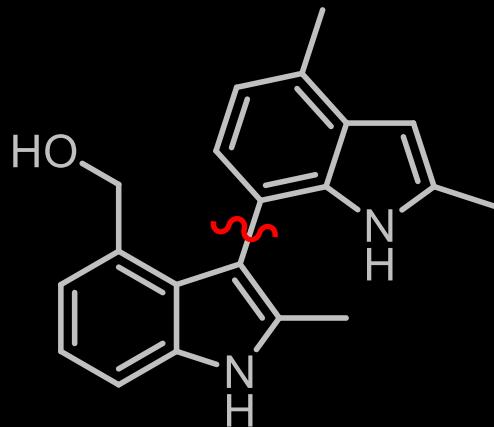
Synthetic Experts

SciFinder

AI^Zynthfinder

ChemPlanner

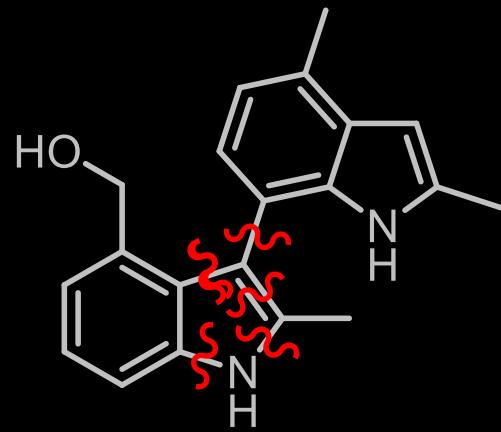
ASKCOS



is that it?

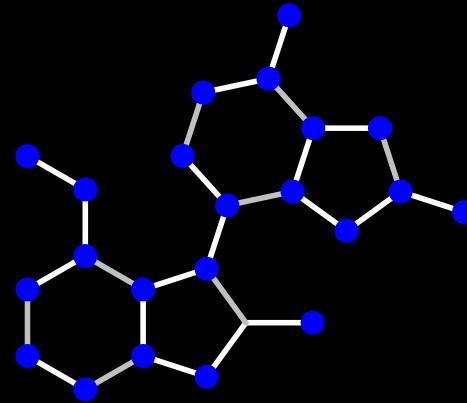
ROUTE SCOUTING

finding transformations



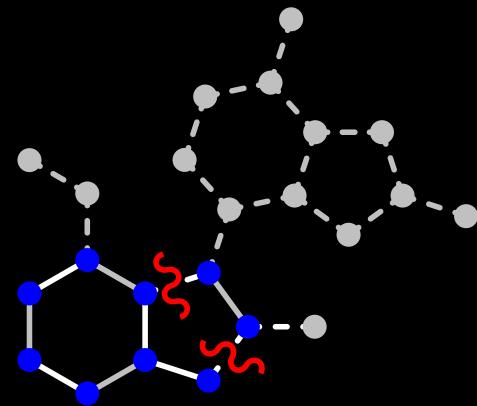
ROUTE SCOUTING – i&os ROUTINE

molecular (sub)graph

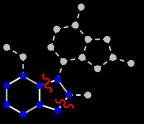


ROUTE SCOUTING – i[&]os ROUTINE

molecular (sub)graph



ROUTE SCOUTING – i&os ROUTINE



for each disconnection

SUGGESTED

$A + B \rightarrow P$

fragment score

R-group score

similarity

transform descriptors

reaction descriptors

REPORTED

$A_1 + B_1 \rightarrow P_1$

$A_2 + B_2 \rightarrow P_2$

$A_3 + B_3 \rightarrow P_3$

.....

$A_j + B_j \rightarrow P_j$

ROUTE SCOUTING – **i&os;** ROUTINE

semi automatic process

SMILES in column one or two. Any additional columns are dropped.

```
:param file_name: (incl path)
:param delimiter: optional
:param header: None, False, True (enforces): row 0. False auto detects line 0 vs 1 if header or not
:return: pandas table w smiles in first column & whatever identifier named as 'ID'
"""
if header is None:
    header = False # due to later analysis

if not header:
    # a form of header auto detection, but only for first line.
    with open(file_name) as file_in:
        first_line = file_in.readline().split()
    first_col = first_line[0]
    # check if more than just one column
    if len(first_line) > 1:
        second_col = first_line[1]
    else:
        second_col = "dummy"
    # finally auto set header true/false
    if is_smiles(first_col) or is_smiles(second_col):
        header = False
    else:
        header = True

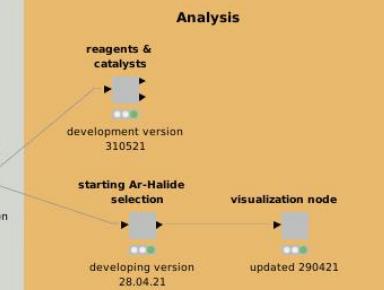
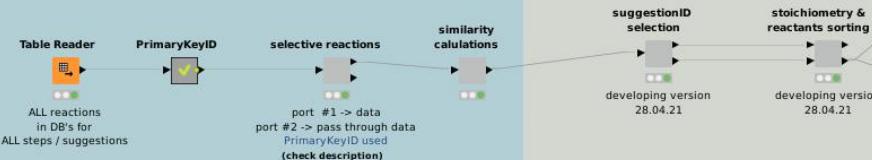
if not header:
    df = pd.read_csv(file_name, sep=delimiter, header=None, dtype=None)
else:
    df = pd.read_csv(file_name, sep=delimiter, header=0, dtype=None)
header = True # revert value for later analysis

# determine smiles or text for cells 0,0 and 0,1
loc00 = is_smiles(df.iloc[0][0]) # smiles or text
if len(df.columns) >= 2:
    # we don't want here other columns. becomes too much.
    df.drop(df.columns[2:], axis=1, inplace=True)
    loc01 = is_smiles(df.iloc[0][1]) # smiles or text

# create artificial name
if len(df.columns) == 1:
    df[1] = "Mol" + df.index.astype(str)
    loc01 = False
    if not loc00:
        # loc00 is not a string, but contains the header name of the smiles
        # N.B.: we don't have headers yet; they will be moved up later, thus overwrite cell 0,1
        df.iloc[0, 1] = "ID"
    if header:
```

DATA INPUT

The table with the project data is (normally) prepared in other workflow.
Once it's done, the data is analysed here. The first nodes are meant to prepare the data and eliminated unwanted reactions (non-selective reactions in terms of several reaction products)



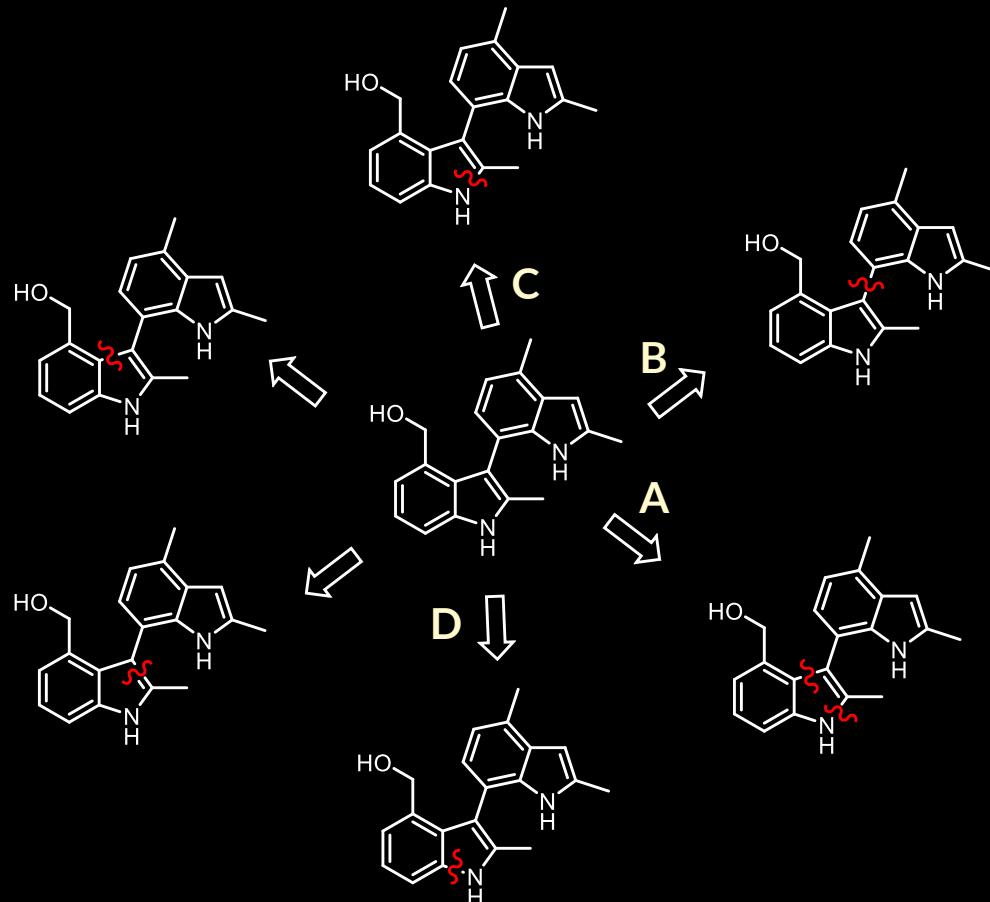
Analysis						
# examples w desired similarity	Max(similarity)	similarity cut-off	Mean(yield)	Unique count(yield)	suggestionID	Search:
123	0.773584902865295	0.4	63.94949494949498	58	2.1	
1179	0.651515126228325	0.4	75.4689655172414	57	2.2	
55	0.7909091114997864	0.4	53.99999999999998	28	2.3	
9	0.44477611780166626	0.4	30	1	2.4	
20	1	0.4	51.5	18	2.5	

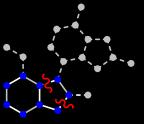
Showing 1 to 5 of 6 entries

Analysis						
comments	desired_product_struct	wrong substitution pattern (P)	# examples w right subst. pattern	Mean(yield)	suggestionID	Search:
key structural motifs present in the suggestion step	<chem>O=C1C=C(C=C1Br)N(C)C(=O)c2ccccc2</chem>	1	547	64.88617886178862	2.1	
key structural motifs present in the suggestion step	<chem>O=C1C=C(C=C1Br)N(C)C(=O)c2ccccc2</chem>	33	1856	67.8727414330219	2.2	
key structural motifs present in the suggestion step	<chem>O=C1C=C(C=C1Br)N(C)C(=O)c2ccccc2</chem>	0	383	54.41250000000016	2.3	
key structural motifs present in the suggestion step	<chem>O=C1C=C(C=C1Br)N(C)C(=O)c2ccccc2</chem>	0	33	43.5	2.4	
key structural motifs present in the suggestion step	<chem>O=C1C=C(C=C1Br)N(C)C(=O)c2ccccc2</chem>	0	20	51.5	2.5	

Showing 1 to 5 of 6 entries

(almost) all synthetic alternatives



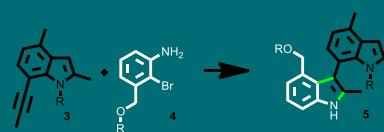


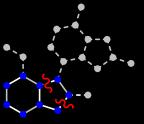
from computer to lab bench

reported
reactions

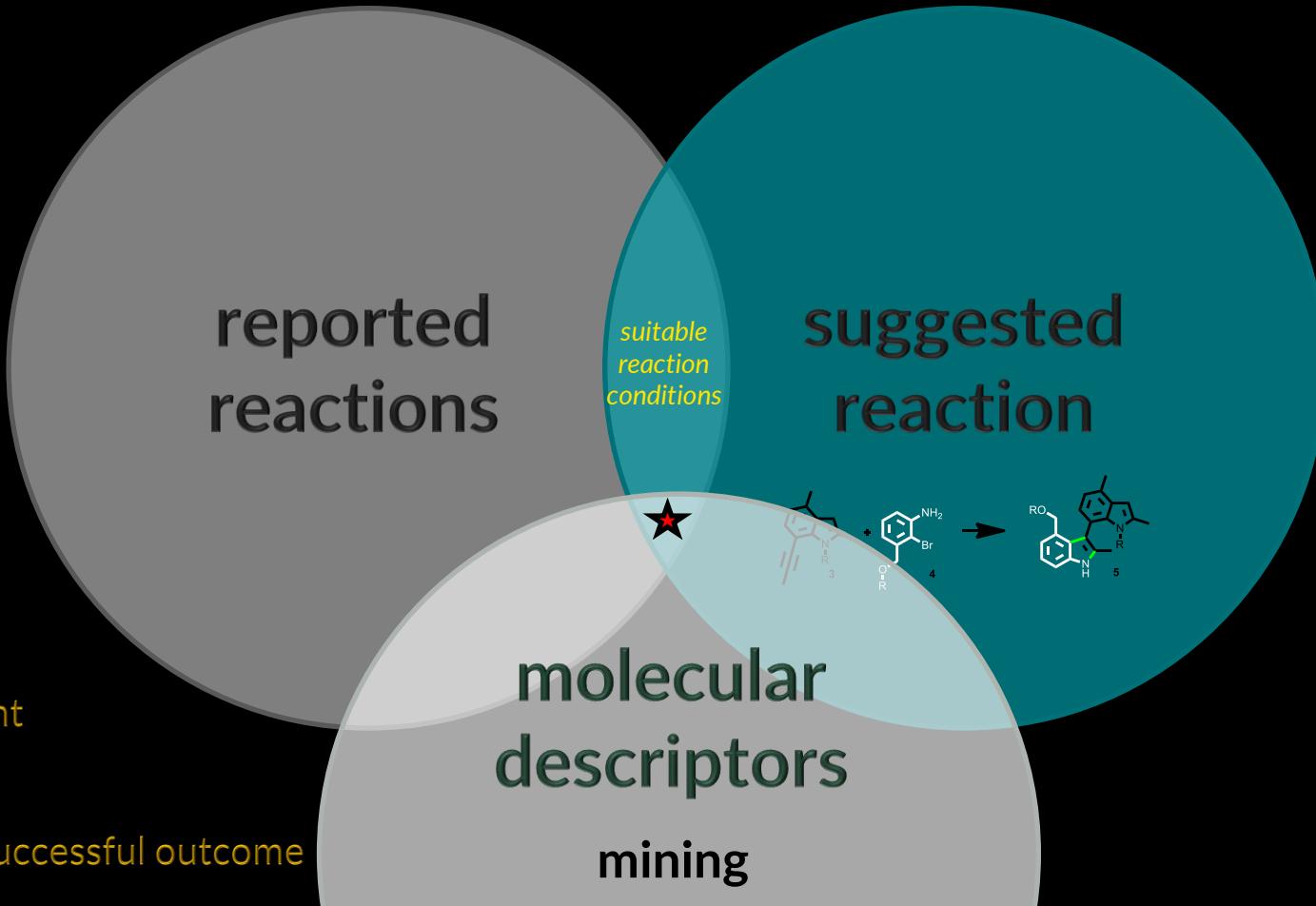
*suitable
reaction
conditions*

suggested
reaction





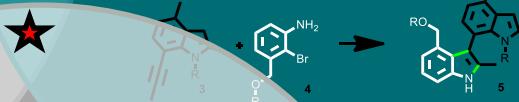
from computer to lab bench



reported reactions

suggested reaction

*suitable
reaction
conditions*

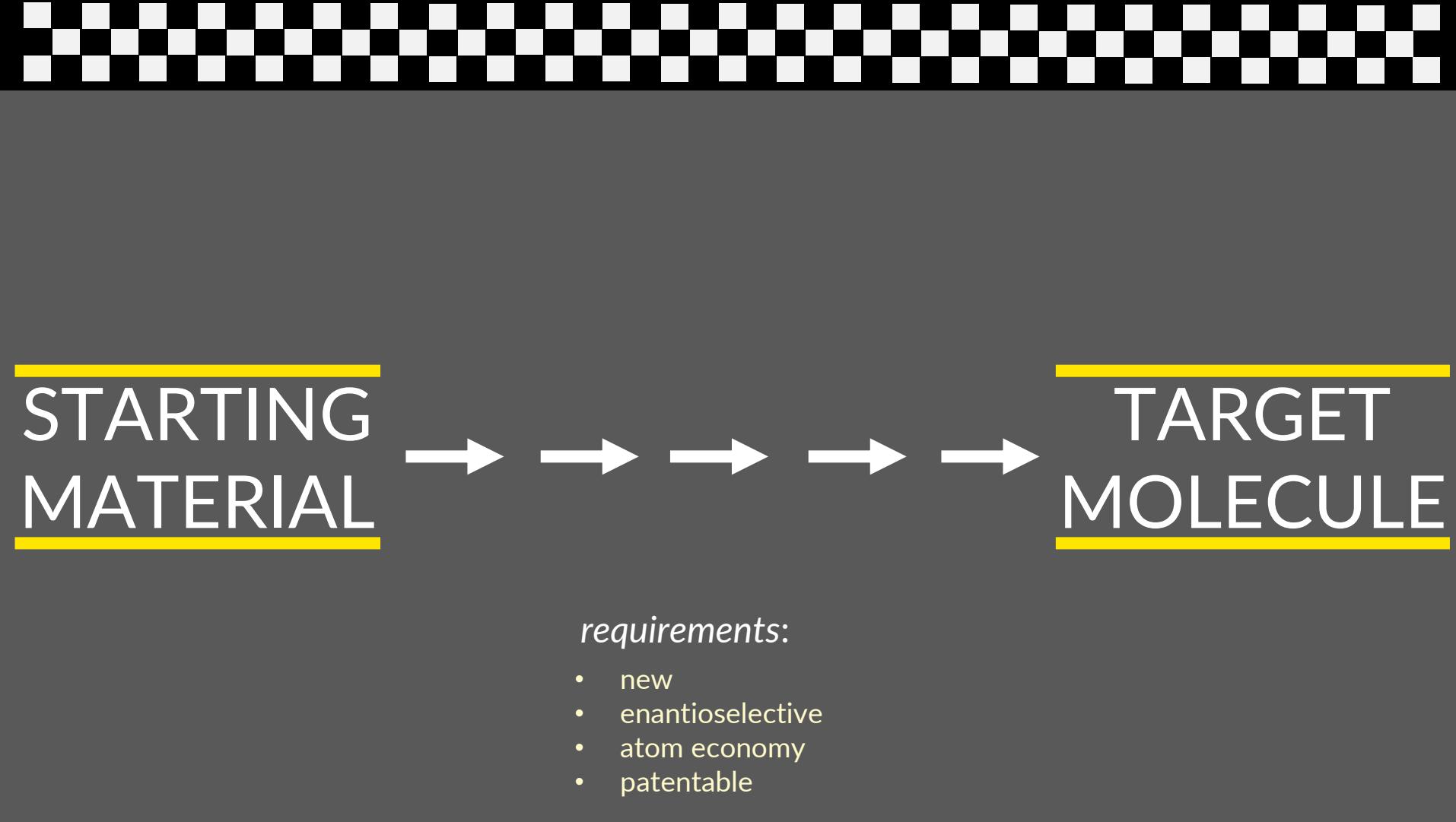


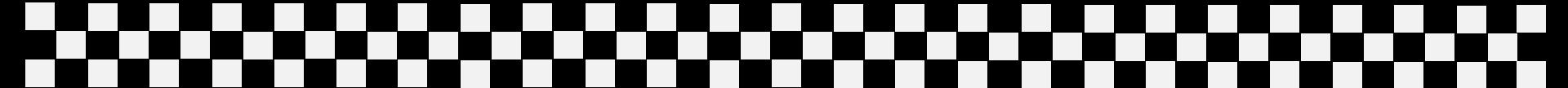
molecular descriptors

mining

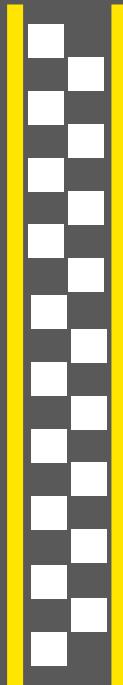
- cost efficient
- less waste
- increased successful outcome

WHY?





STARTING MATERIAL

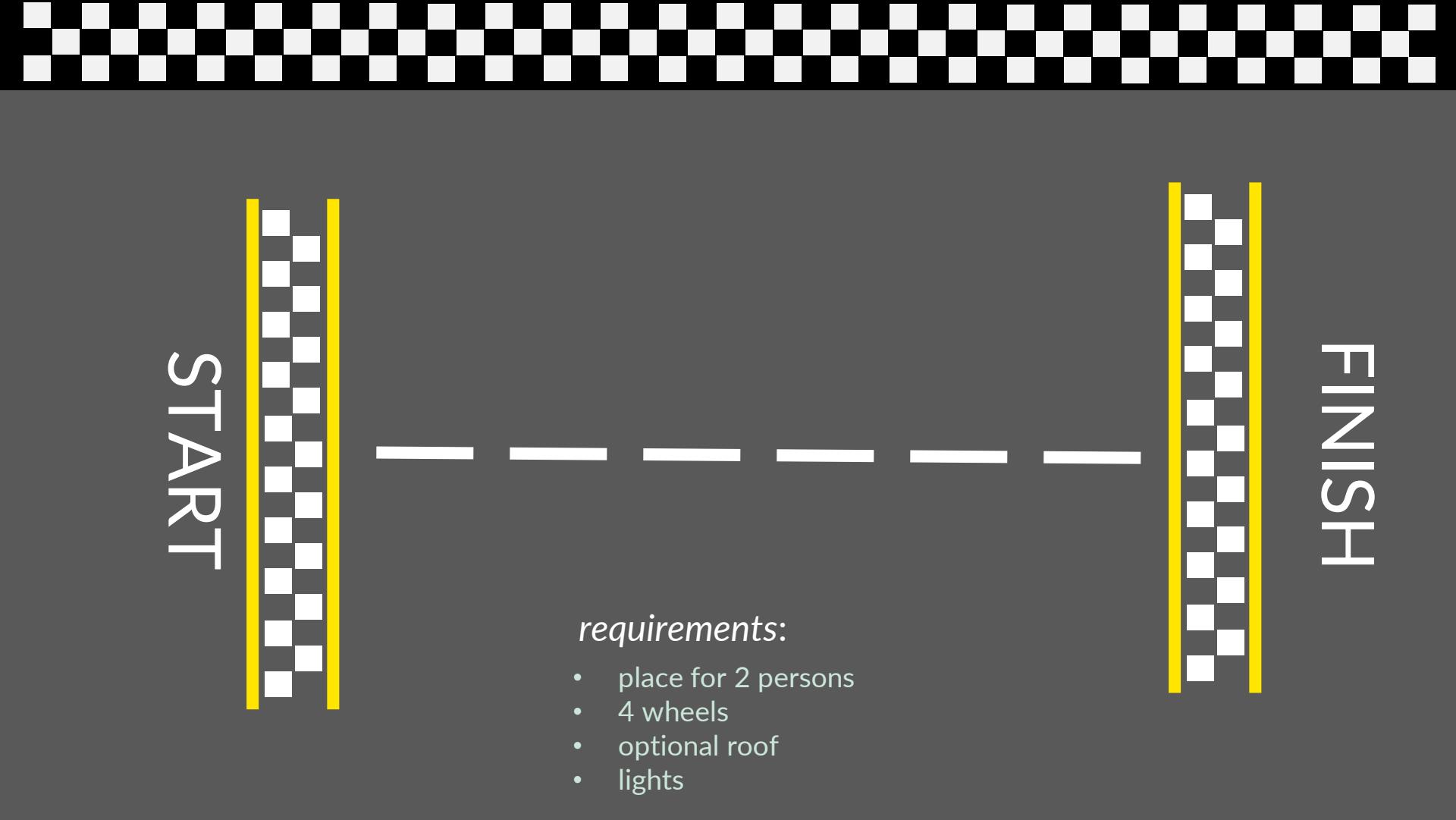


TARGET MOLECULE



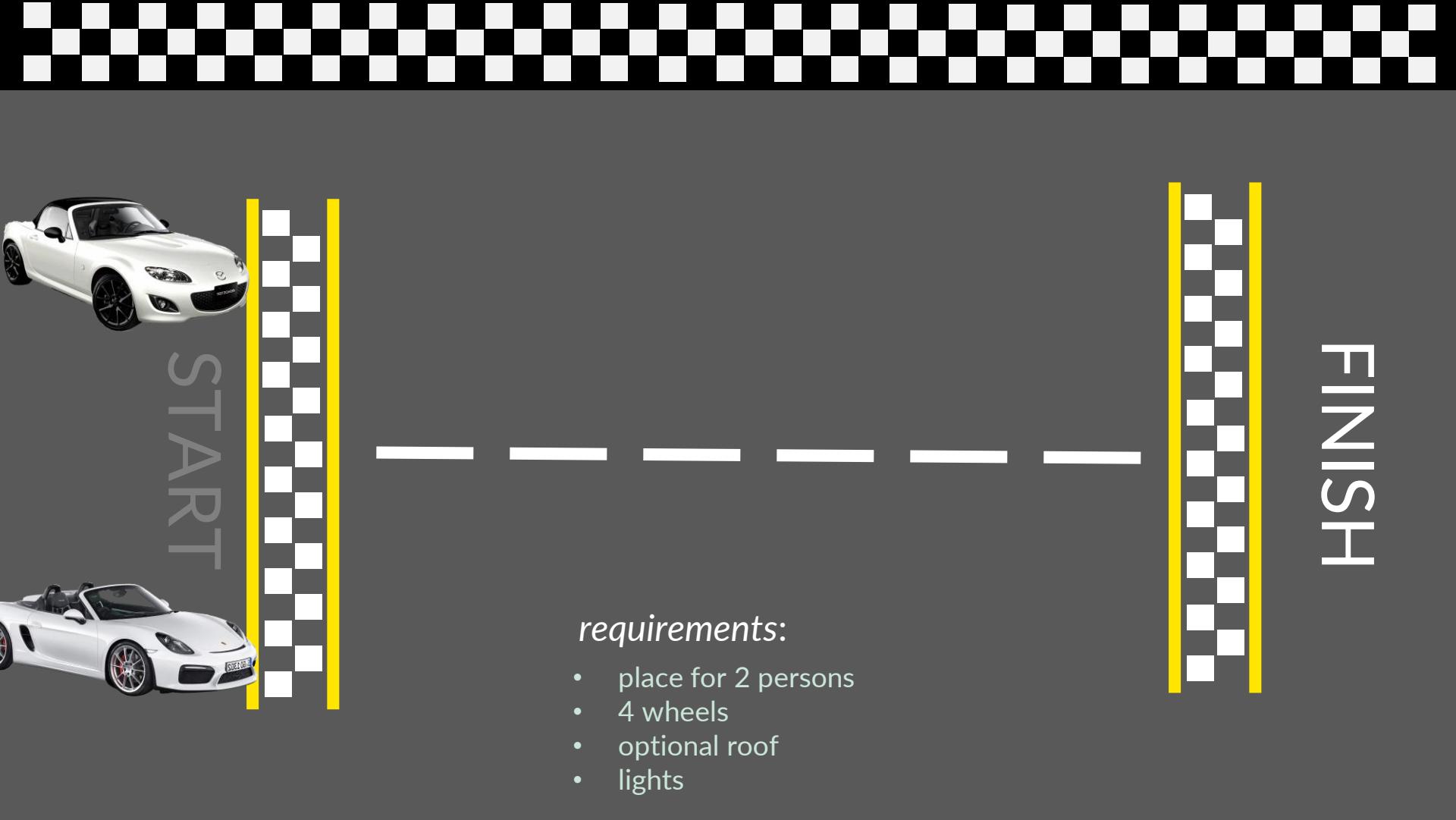
requirements:

- new
- enantioselective
- atom economy
- patentable



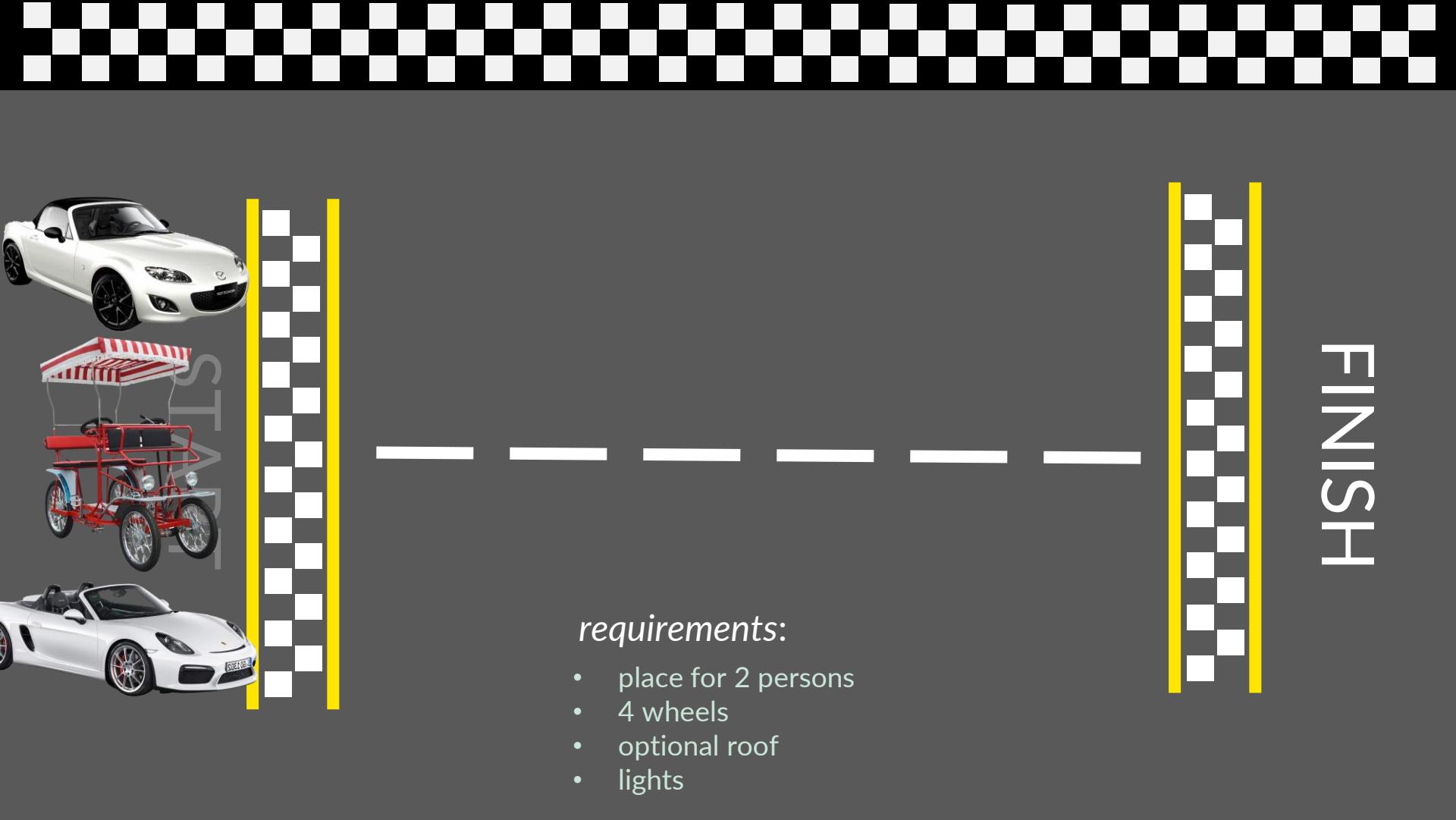
requirements:

- place for 2 persons
- 4 wheels
- optional roof
- lights



requirements:

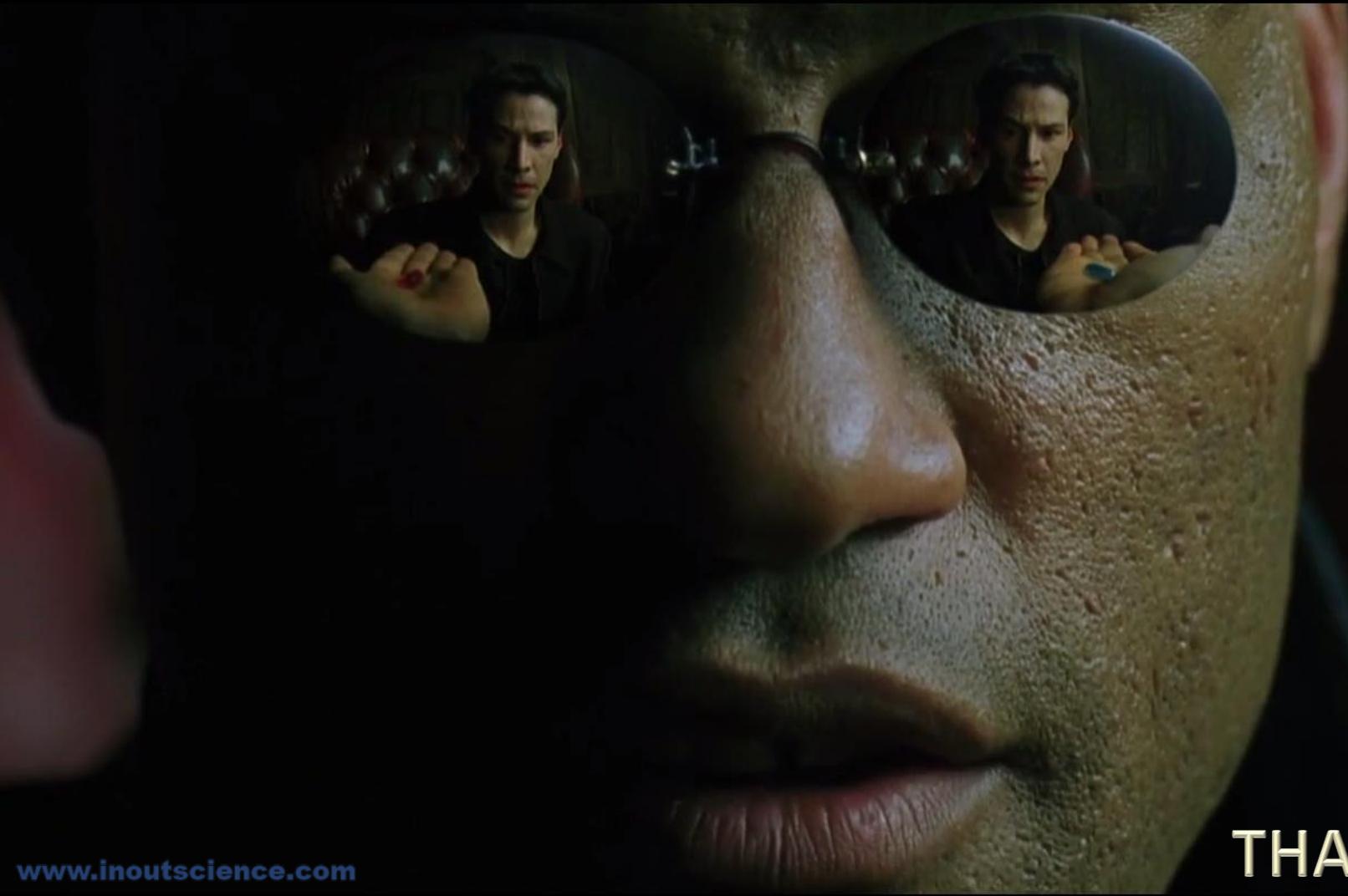
- place for 2 persons
- 4 wheels
- optional roof
- lights



requirements:

- place for 2 persons
- 4 wheels
- optional roof
- lights

FINISH



THANK YOU