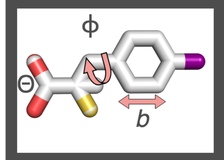
# Learned chemical perception of force field typing rules using Monte Carlo sampling

Camila Zanette, Caitlin C. Bannan, Christopher I. Bayly, Josh Fass, Michael K. Gilson, Michael R. Shirts, John D. Chodera, David L. Mobley

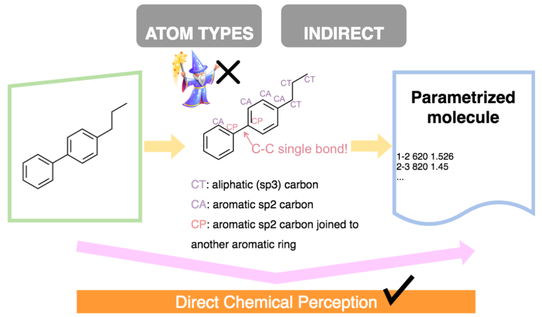
University of California Irvine, Memorial Sloan Kettering Cancer Center, OpenEye Scientific Software Inc., University of California San Diego, University of Colorado Boulder

## BACKGROUND

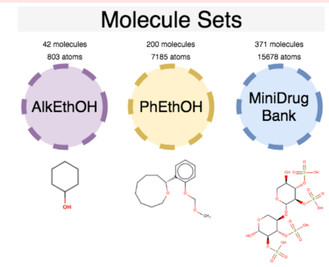
- FORCE FIELDS (FF) GIVE THE ENERGY AND FORCES OF AN ATOMISTIC SYSTEM AS A FUNCTION OF THE COORDINATES.
- WE USE FF TO RUN MD OR MC SIMULATIONS.
- CONSIDERABLE HUMAN EXPERTISE IS REQUIRED FOR FF DEVELOPMENT.
- CHEMICAL PERCEPTION THROUGH ATOM TYPES (INDIRECT) CAN RELY IN EXPERT HUMAN INTUITION AND HAND-TWEAKING, CAN RESULT IN A POOR DATA STRUCTURE, AND CAN LEAD TO MANY REDUNDANT PARAMETERS.




WE DEVELOPED TWO TOOLS → SMARTY → ATOM TYPES; SMIRKY → FRAGMENTS




## INPUT

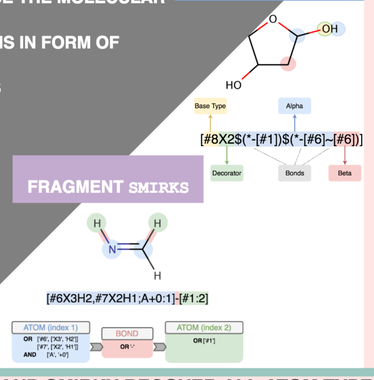



**GOAL: BE ABLE TO LEARN CHEMICAL PERCEPTION OF MOLECULES WHILE SAMPLING OVER ATOM TYPES OR FRAGMENTS**

## METHODS

SMARTS AND SMIRKS LANGUAGES ARE USED TO DESCRIBE THE MOLECULAR PATTERNS IN FORM OF STRINGS

**ATOM TYPE SMARTS**

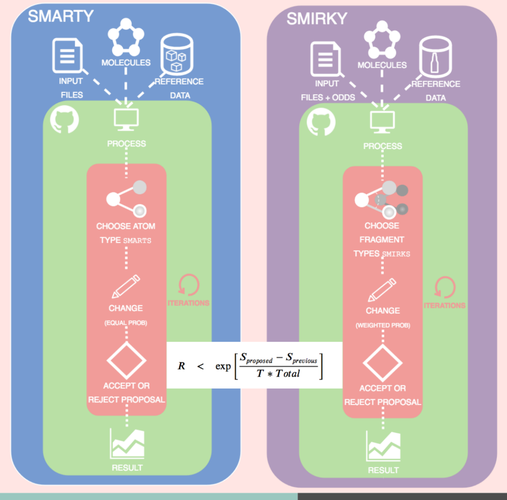



**FRAGMENT SMIRKS**



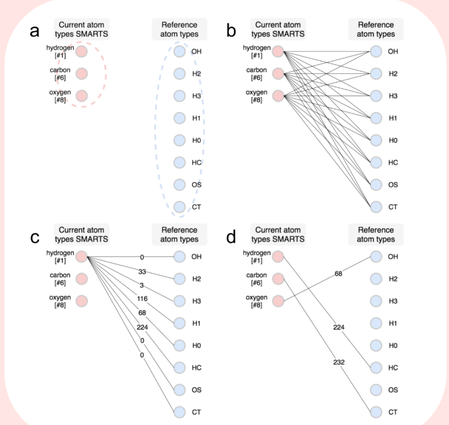

## SMARTY AND SMIRKY REPLACE HAND PICKED ATOM TYPES USING MONTE CARLO SAMPLING

### WORKFLOW




$$R < \exp\left[\frac{S_{proposed} - S_{previous}}{T * Total}\right]$$

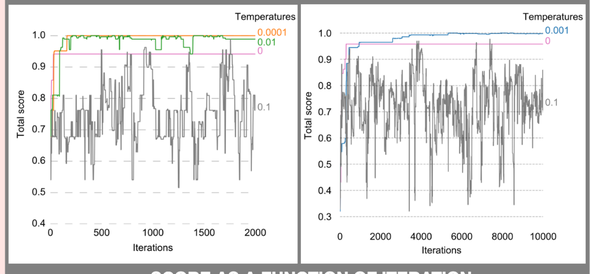
### PROPOSED ATOM TYPES OR FRAGMENTS ARE SCORED USING A BIPARTITE GRAPH




## RESULTS

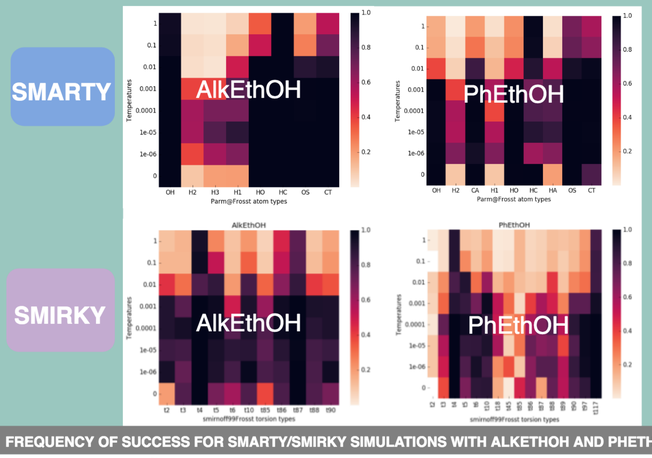
**SMARTY** **SMIRKY**

**100 % SUCCESS ON *AlkEthOH* AND *PhEthOH* SETS**




**SCORE AS A FUNCTION OF ITERATION**

### SMARTY AND SMIRKY RECOVER ALL ATOM TYPES AND FRAGMENTS OF *AlkEthOH* AND *PhEthOH*




**FREQUENCY OF SUCCESS FOR SMARTY/SMIRKY SIMULATIONS WITH ALKETHOH AND PHETHOH**

### *MiniDrugBank* IS SIGNIFICANTLY MORE CHALLENGING FOR SMARTY AND SMIRKY

| | AlkEthOH | PhEthOH | MiniDrugBank |
|---|---|---|---|
| SMARTY * | 100% | 100% | 89.2% |

| SMIRKY ** | BOND | ANGLE | TORSION | NONBONDED |
|---|---|---|---|---|
| AlkEthOH | 5/5 | 4/4 | 11/11 | 8/8 |
| PhEthOH | 10/10 | 6/6 | 16/16 | 9/9 |
| MiniDrugBank | 68/73 | 32/34 | 117/136 | 25/26 |

*PERCENTAGEM OF ATOM TYPES FOUND.
**NUMBER OF FRAGMENTS FOUND OVER THE TOTAL NUMBER OF FRAGMENTS.

## CONCLUSION

THIS STUDY DEMONSTRATES PROMISING RESULTS WHERE WE CAN REPLACE THE HUMAN EXPERT WITH AN AUTOMATED MACHINERY FOR SAMPLING DIRECT CHEMICAL PERCEPTION.

## ACKNOWLEDGEMENTS

- National Science Foundation (CHE 1352608)
- National Institutes of Health (1R01GM108889-01)
- Computing support from the UCI GreenPlanet cluster, supported in part by NSF Grant CHE-0840513
- Brazilian Science without Borders scholarship (Capes - BEX 1865612-9)

openforcefield.org — mobleylab.org
**czanette@uci.edu**