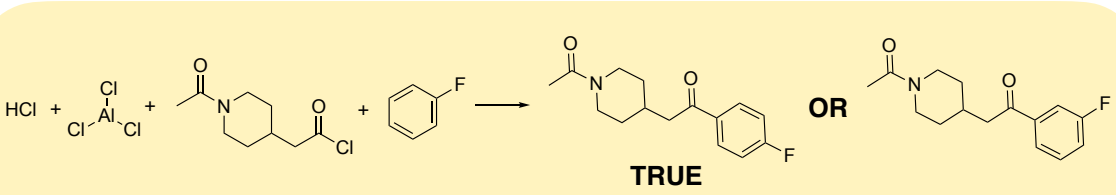
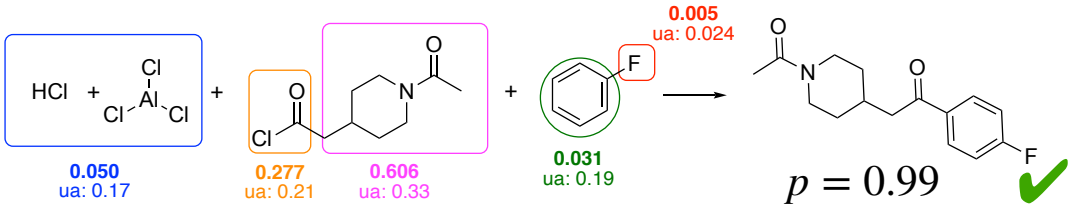
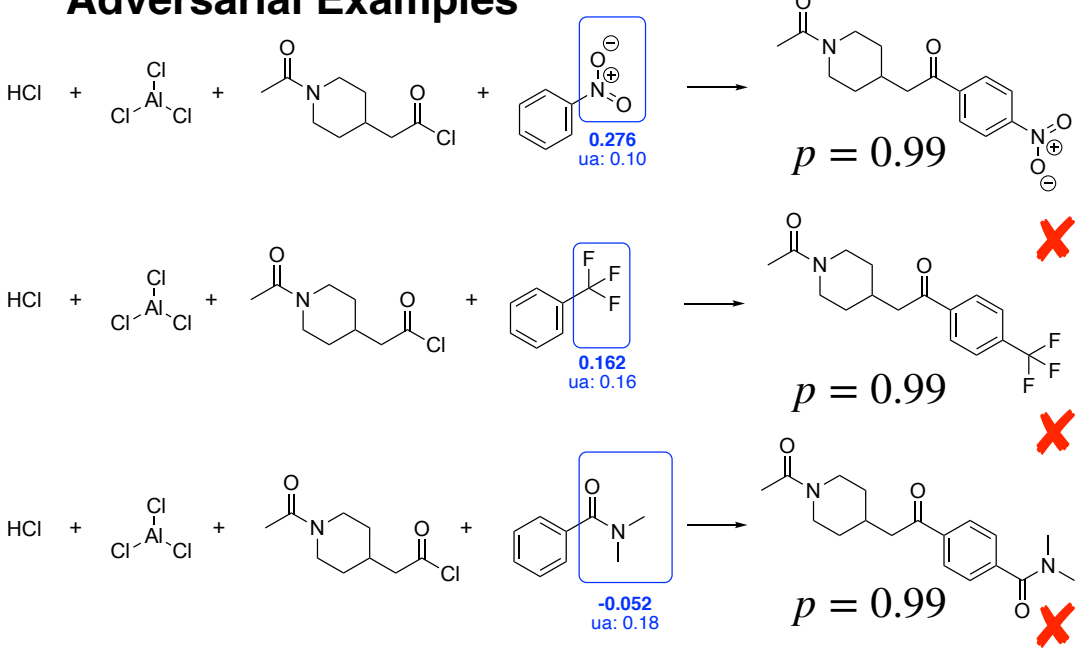
(a) **Input Reaction**



Model Top-1 Prediction + Input Attribution



ua = uniform attribution

(b) **Adversarial Examples**