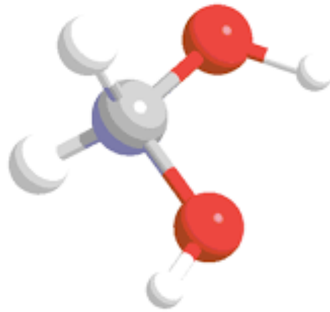


Inter-molecular Graph Pathfinder

Graph Mining Project Proposal

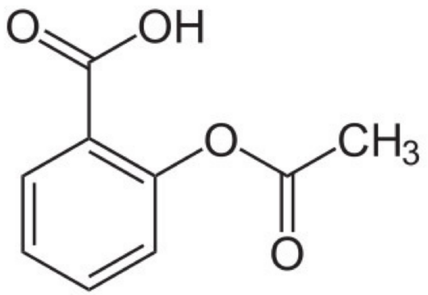
{ Interactive Graph Mining }

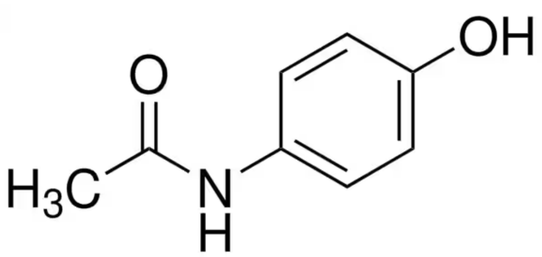
성균관대학교 소프트웨어학과 김산

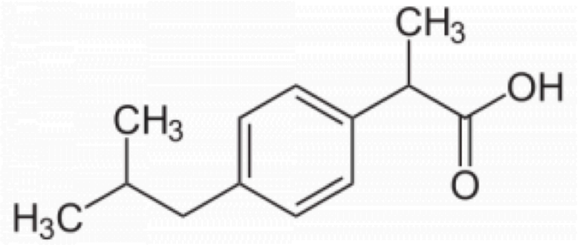
2023. 10. 11. 수요일.

Introduction

Pharmaceutics :: 제약

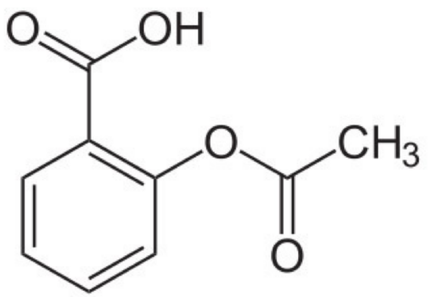




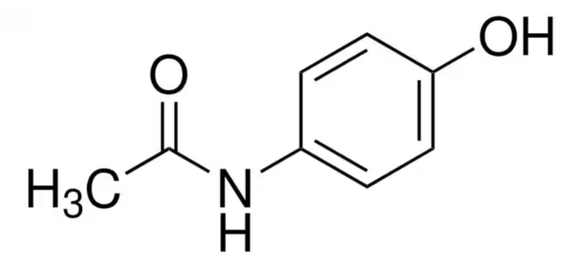


Introduction

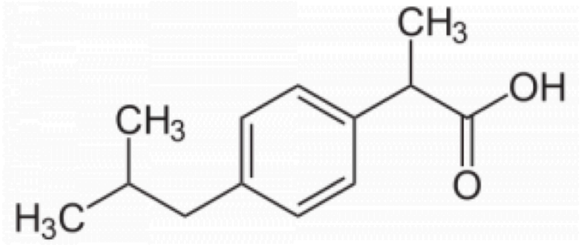
Pharmaceutics :: 제약



Acetylsalicylic acid



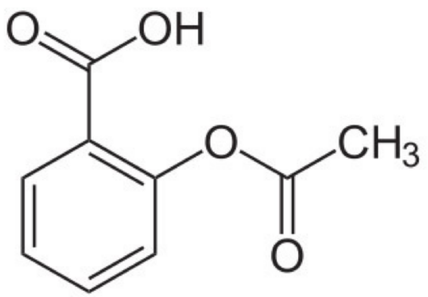
Acetaminophen



Ibuprofen

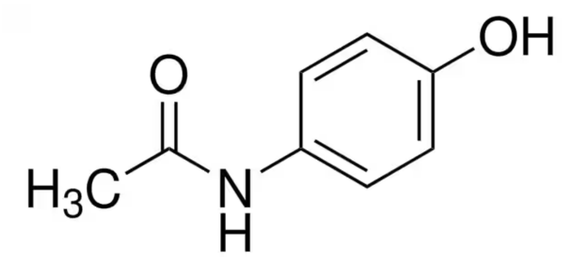
Introduction

Pharmaceutics :: 제약



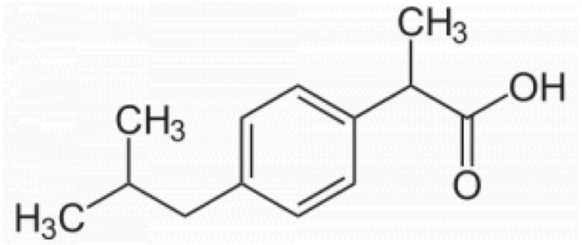
Acetylsalicylic acid

아세틸살리실산



Acetaminophen

아세트아미노펜

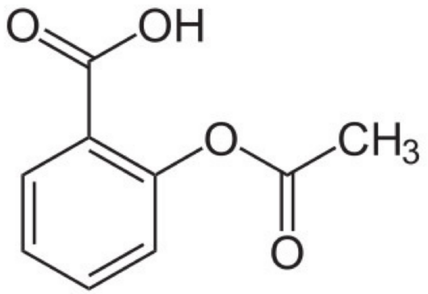


Ibuprofen

이부프로펜

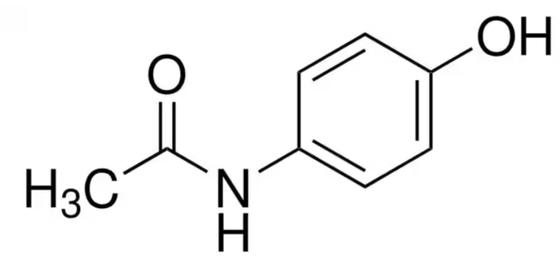
Introduction

Pharmaceutics :: 제약



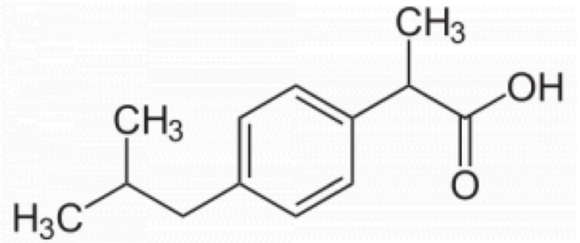
Acetylsalicylic acid

아세틸살리실산



Acetaminophen

아세트아미노펜

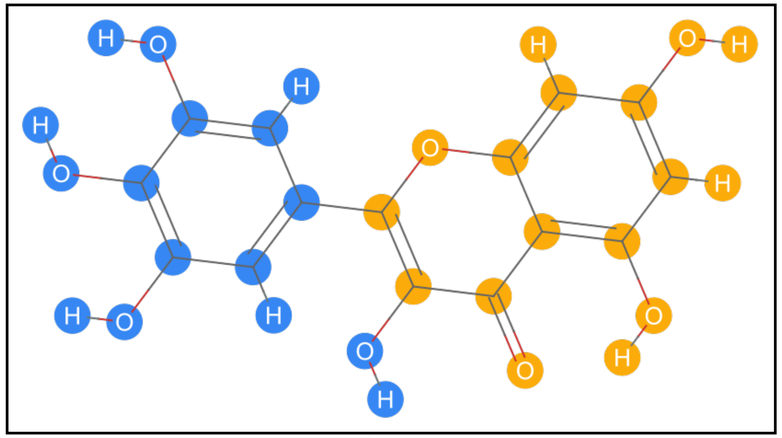


Ibuprofen

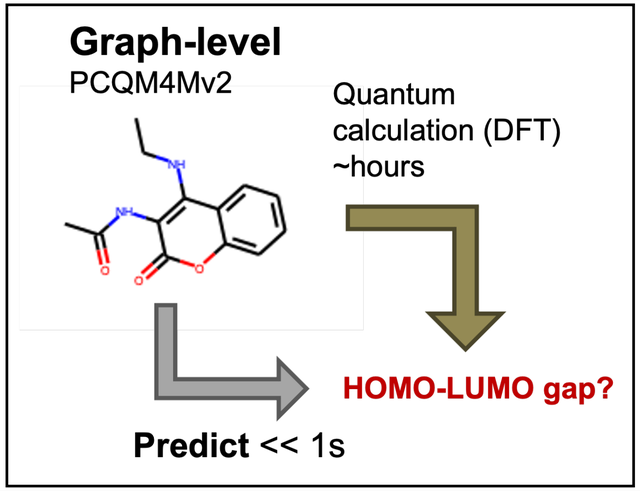
이부프로펜

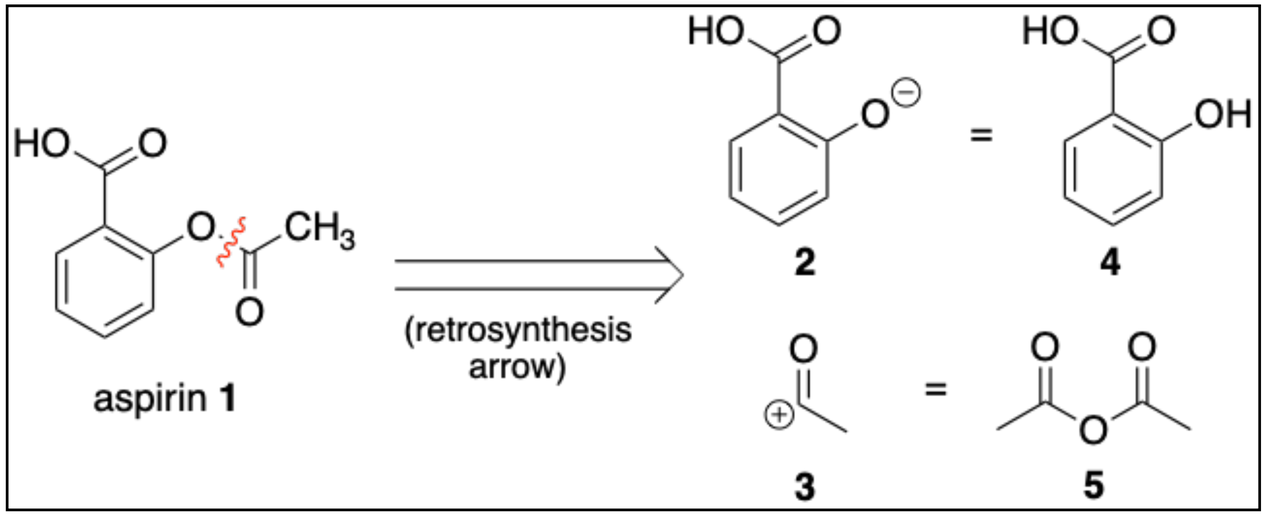
Motivation

어떻게 chemical graphlet들을 찾을 수 있을까?



- **(Intra) Molecular graph**
  - Node = atom + atomic feature
  - Edge = bond + bond feature
- **작용기 (Functional group)**
  - Graphlets, Inter-molecular feature
- **역합성 (Retrosynthesis)**
  - 제약, 화학공학





Dataset

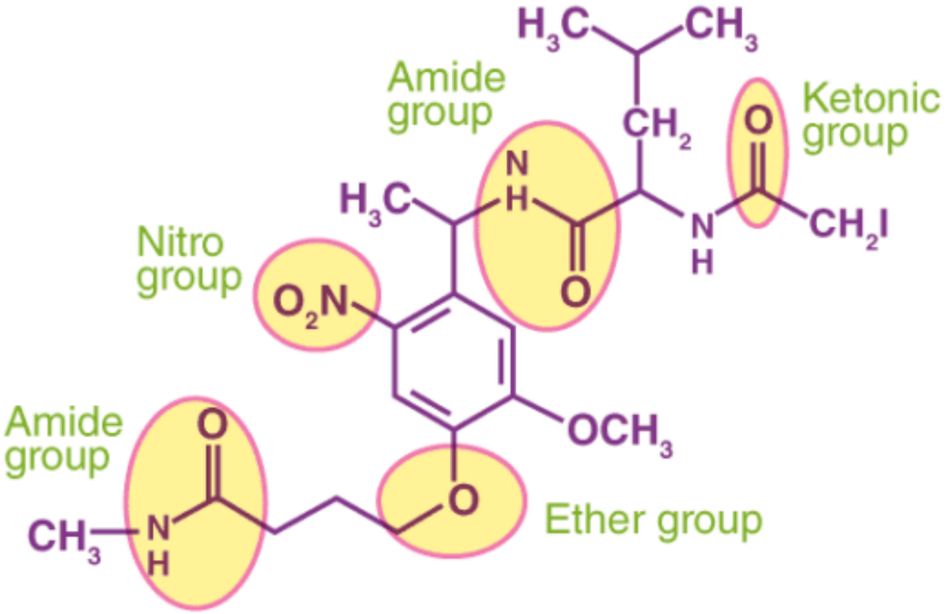
QM7 dataset → Inter-molecular graph

- **Quantum Machines 9 (QM9) dataset**
  - Nodes
    - Atoms
  - Edges
    - Molecular bonds
      (Valence Shell Electron Pair Repulsion theory based)
  - # of graphs
    - 7165 graphs(= molecules)
    - MAX 23 nodes(= atoms) per graph
      - MAX 7 heavy atoms per graph

- **Inter-molecular graph**
  - Nodes
    - Molecules
  - Edges
    - **Subgraph relations (Δatom = 1)**
  - 1 large graph
    - 7165 nodes(= molecules)

Purpose

Reveal undefined functional groups



- **작용기 확인하기**
  - Functional group은 화학적으로 안정하면서도
  - 자주 나타나기 때문에 molecular graph의 edge feature를 수집해 발견할 수 있음
- **작용기를 분류하기 :: Inter-molecular feature**
  - 기존에 알려진 작용기들과 비슷한 빈도로 등장하면서
  - 특별한 기능으로 명명되지 않았던 작용기를 계층적으로 분류해 **inter-molecular** feature로 사용할 수 있음

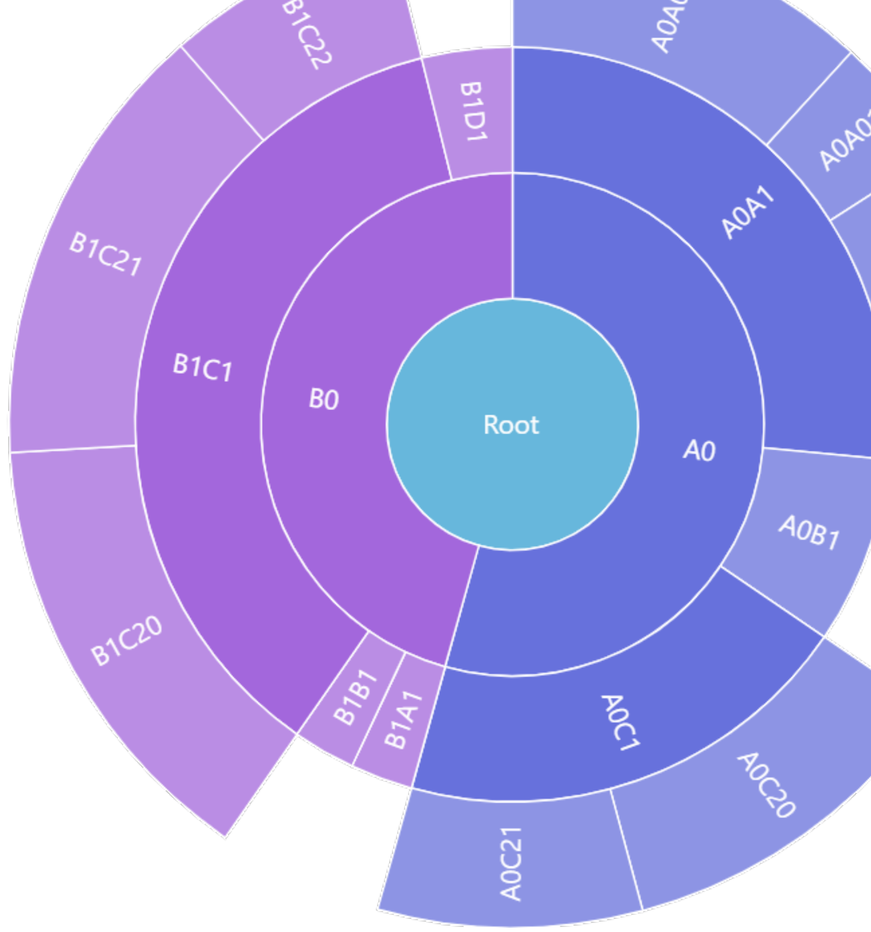
Method

Sample pathways from Eigen-central nodes

- **Inter-molecular graph의 Eigenvector-centrality 계산**
  - 주요 backbone 분자구조 indentify
  - 모든 path를 계산하는건 computational/physical 문제가 있기 때문에 중심노드 선정
- **Path 정보 수집**
  - Eigenvector-center로 부터 커지는 방향으로 원자들을 붙여나가면서
    지나가는 edge 정보 수집
- **Path dataset 분석**
  - Hierarchical clustering을 통해 자주 등장하는 path를 hierarchical 분류

Expected Result & Conclusion

Hierarchical clustering of functional groups



- **예상 결과**
  - **Inter-molecular graph**를 정의하고 구축
  - Chemical path를 hierarchical clustering
    - 기존에 알려진 작용기와 비교 분석
- **의의**
  - **Inter-molecular** feature를 **graph** mining 기법으로 발견하려는 시도
  - Path를 역으로 밟아 retrosynthesis software에 응용 가능
  - Learning project에 사용 가능