



# DB- Chemical Database for Drug Design -

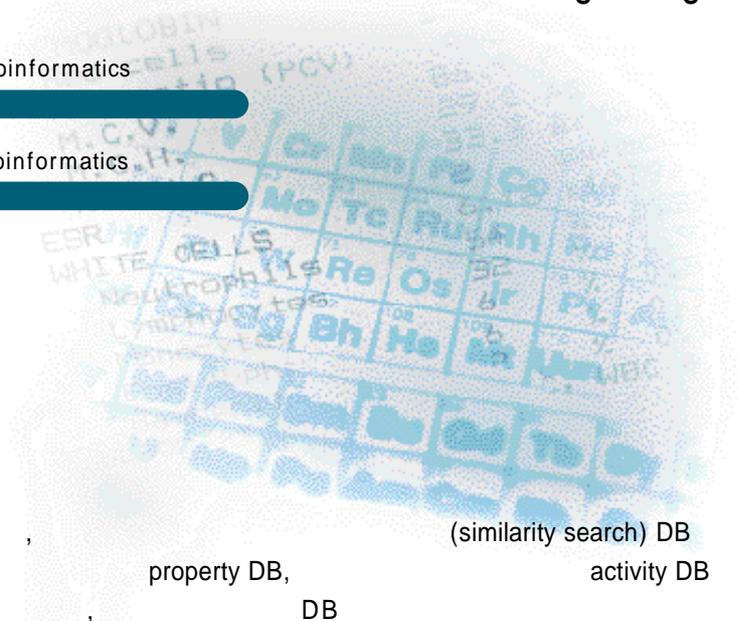
Chem & Bioinformatics

leesk@qspr.pe.kr

Chem & Bioinformatics

ktno@camd.ssu.ac.kr





DB, reaction DB, DB, 가 DB, (similarity search) DB, activity DB, property DB, DB

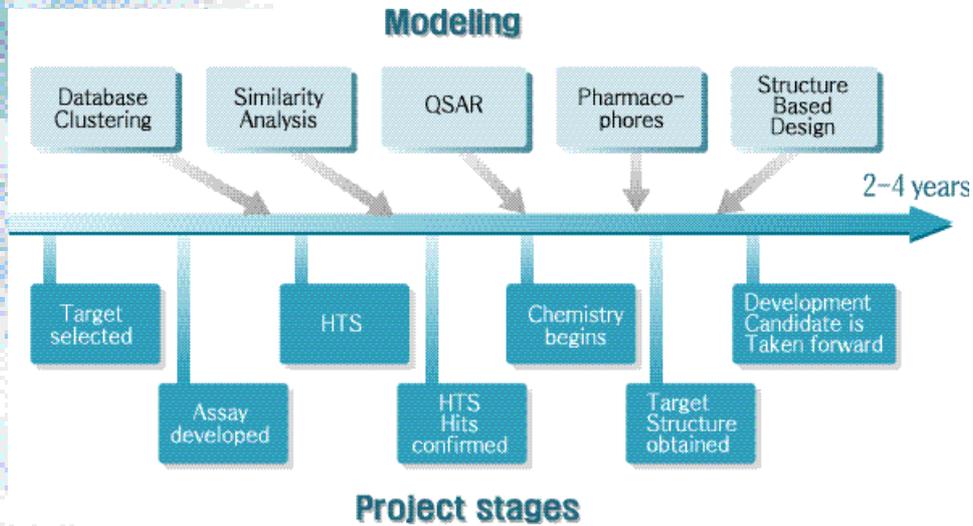
I. DB clustering, 가 QSAR

(lead compound), phamacophore query, DB, (structure based design), (lead compound), (DB), DB

(high throughput screening, HTS), (combinatorial chemistry), (compound library), 가 DB, (similarity search) DB, reaction DB, property DB, activity DB

DB, DB, 가

# 과학기술종합정보시스템

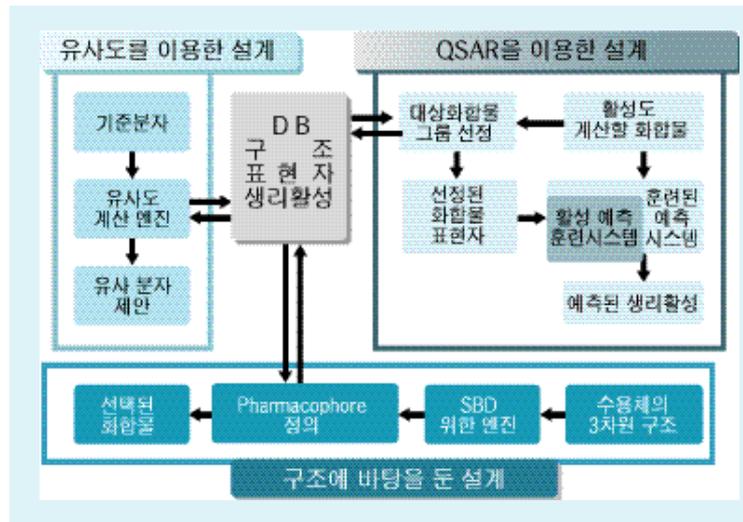


< 1 >

DB   2   pharmacophore   DB

가   102 106   가   가

(drug-like) DB   가



< 2 >

X-ray   NMR   3   가   가

SBD   가

pharmacophore   , DB   3)



< 3> DB (<http://chemdb.kordic.re.kr>)

DB가  
(<http://chemdb.kordic.re.kr>) 3).  
( )  
2000  
(<http://camd.ssu.ac.kr>) ) (<http://www.cnbr.co.kr>) ) (<http://www.ktit.com>)  
DB  
web ( )  
)  
3  
DB  
가 DB

II.

1. DB

DB

< 1>

index

21

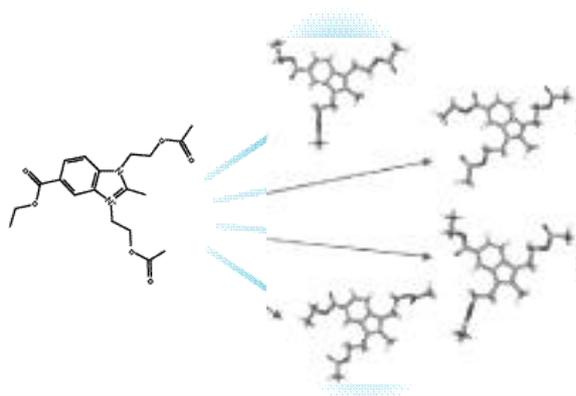
2

< 2>.

< 1> NGI

index	ABX (advanced biochemical compound) catalog, Acros, Aldrich, Building Blocks catalog 2000, CRC handbook of chemistry and physics, DAVOS product catalog, Fluka chemical, GFS chemicals, ICN biomedical catalog, INCOFINE Chemical catalog, Junsei chemicals, Kanto chemicals, Lancaster, LOBA organics 1999, Merck catalog, Research organics, TCI, The Merck Index, The millennium catalog (AccuStandard), Tocris, Wako - 21 (300,000) Journal of Organic Chemistry (1969-1998) Journal of Medical Chemistry (1992-1997) Journal of Chemical Information and Computer Science (1992-2000) Journal of the American Chemical Society (1991-2000) Tetrahedron Letter (1999-2000) - 5 (310,000) DB National Cancer Institute Database (240,000 )
-------	---





< 6> 2      3      가

가

(energy barrier)  
conformational search

가 (dihedral

angle) random search <sup>10)</sup>

conformation search

가 ±20° 가 conformer

4    3

< 3>      가

1	atom count	(2D )
2	bond count	(2D )
3	connectivity index 0	(2D )
4	connectivity index 1	(2D )
5	connectivity index 2	(2D )
6	valence connectivity index 0	(2D )
7	valence connectivity index 1	(2D )
8	valence connectivity index 2	(2D )
9	calculated logP	(2D )
10	No. hydrogen bond acceptors	(2D )
11	No. hydrogen bond donors	(2D )
12	Molar refractivity	(2D )
13	Moment of Inertia	(3D )
14	Molecular surface	(3D )
15	Molecular volume	(3D )
16	HOMO Energy	Highest Occupied Molecular Orbital(HOMO) energy (3D )
17	LUMO Energy	Lowest Unoccupied Molecular Orbital(LUMO) energy (3D )
18	Dipole moment	(3D )

corina <sup>11)</sup>      3

.< 6>      3

semi-empirical method  
(single point calculation)      HOMO,  
LUMO      (dipole moment)

(moment of inertia),

가      3

3.

DB      4

2D      substructure

2D

3D

7)

scalar

common name, IUPAC name, CAS

number

(query structure)

가

가

substructure

가 < 8> ,

가

# 과학기술종합정보시스템

**일반정보를 이용한 검색**

대개 일수를 기준으로 하되, 분자번호를 이용한 검색의 경우에는 분자번호(MW), 분자  
중량, 분자부피 등을 이용한 검색이 가능합니다.

특수한 검색 : 이점이나 화학식으로부터 검색할 원하는 그룹을 선택하십시오.

Common Name

MPAC Name

CAS Registry Number

화학식 예 : C<sub>10</sub>H<sub>16</sub>O<sub>2</sub>N<sub>2</sub>

---

Keyword 검색 : 다음의 옵션과 일들을 이용하여 검색을 할 수 있습니다.  
(단, 분자번호를 이용한 검색은 불가능합니다.)

분자량 범위  From :  To :

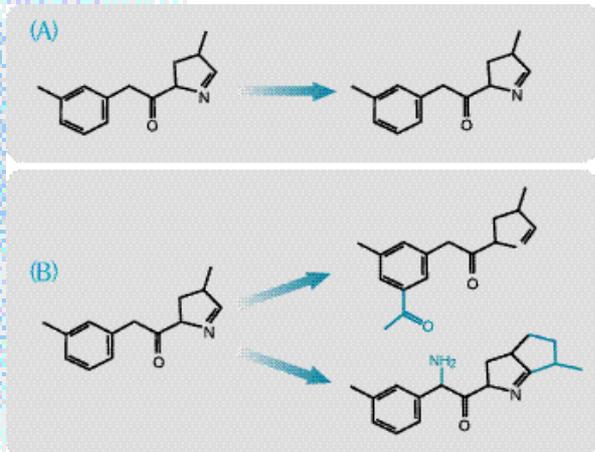
분자 표정식  From :  To :

분자 번호  From :  To :

< 7 >

substructure

labeled graph subgraph  
isomorphisms  
가 Ullman 가  
가 <sup>12)</sup>



< 8 > (A) (B) substructure

Ullman  
substructure searching system

Ullman

<sup>13)</sup>,

C + +

가 (<= 15

)

false negative

가

Subgraph Isomorphism

(A), (B), (C), (O),  
(N), (S), (P) 가

Substructure

substructure

Database

가

compact  
mol file connectivity

table

< 4 >

DB

substructure

. 3

false negative hits가  
(SciFinder

가  
CAS



< 4> DB substructure

3

	1	2	3	DB	
	494	485	485	NCI DB 246,182	web service
SciFinder(CAS) <sup>1)</sup>	330	153	31,336	CAS DB 28,000,000	client program
NCI Database browser <sup>2)</sup>	3	379	451	NCI DB 246,182	web service
ChemFinder <sup>3)</sup> (Cambridge Soft.)	0	0	56(463 <sup>6)</sup> )	NCI DB 246,182	web & client
ISIS/host(MDL) <sup>4)</sup>	0	0	19(206 <sup>6)</sup> )	NCI DB 123,219	client program
NCI 3D <sup>5)</sup> (Chem Web)	0	0	19(206 <sup>6)</sup> )	NCI DB 123,219	web service

1

2

3

1) Chemical Abstract Service, URL <http://www.cas.org>  
 2) NCI Database browser, URL <http://cactus.cit.nih.gov/ncidb2>  
 3) Cambridge soft., URL <http://www.chemfinder.com>  
 4) MDL Information system Inc., URL <http://www.mdli.com>  
 5) ChemWeb.com, URL <http://www.chemweb.com>  
 6)

test

DB  
가

DB negative hits가 false 9

benzene ring

가 chain

prefilter  
가

test

mol file  
topology(  
ring

)

4.

DB

MDL mol file

DB

mol file

DB

Randic and Kier & Hall<sup>4)</sup>  
(1-2)

$i \quad j \quad (i \quad j)$

Randic connectivity index, (6)

Kier & Hall valence connectivity index가

$Z_i \quad i$

$Z_i \quad v$

< 9> False negative hits가

28,000,000

test

test

가

).

가,  $H_i$   $i$

$$^nX = \sum_{\text{(pairs of length } n)}^{N_{SB}} (i_1 \dots i_{n+1})^{1/2} \dots \dots (1)$$

$$i = \frac{(Z_i^v - H_i)}{(Z_i - Z_i^v - 1)} \dots \dots (2)$$

substructure

Higo & Go

Waals, solvent accessible

Van der Waals

Van der Waals

accessible

, solvent 가

. Higo & Go

( 2 )

external surface internal,  
surface

internal surface

5. DB

DB  
30

5

DB

DB

< 5 >

DB

DB

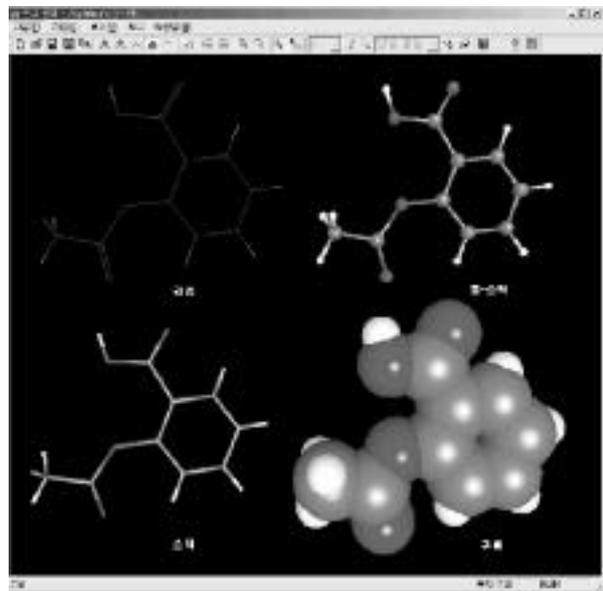
DB		Data	
KISTI	DB	10 ( , , )	2D 850,000
		2D	3D 350,000
		10	
		3D ( conformer )	<2004 >
	6 3D	2D 1,500,000	3D 2,500,000
MDL	ACD 3D	2D	2D 278,093
		가 , , CAS number	3D 227,272
	MDDR 3D	3D ( conformer )	
		2D biological testing	2D 113,842
CMC 3D	3D ( conformer )		
	2D	2D 7,936	
	drug logP, pKa	3D 7,936	
CambridgeSoft.	ACX ACX-screening가 ,	2D	2D 850,000
CAS	SciFinder	2D (fulltext)	2D 28,000,000
Derwent	DDF	2D	2D 1,600,000



MDL ACD 3D DB  
 DB  
 . Cambridge  
 ACX pc  
 가 . MDDR  
 , CMC DB  
 , SciFinder  
 fulltext  
 , Derwent  
 DB  
 . DB  
 1000 가 use가  
 DB  
 , DB  
 . 3  
 3 conformer  
 2  
 .  
 substructure  
 , 가 , DB  
 가 .  
 가 . 가  
 5 ,  
 2D  
 3D 2004  
 , 가 DB  
 .  
 6. Molview

( 10). Source code visual C++  
 opengl library ,

가 , 3



< 10> MolView

가가 DB  
 MDL mol file cerius2  
 msi file,  
 mvl file , MDL mol file  
 mvl file 가 .  
 (line)  
 (stick), - (ball-stick), (Space filling  
 Corey-Pauling-Kolton(CPK) model)

가 가 .  
 x, y, z

Molview v1.0

molview

III. 가상 분자 모델링 및 가상 스크리닝

가상 분자 모델링은 분자의 구조와 특성을 컴퓨터를 이용하여 모사하는 기술로, 분자 모델링 소프트웨어를 사용하여 분자의 구조를 생성하고, 분자의 특성을 예측하는 데 사용된다. 가상 스크리닝은 분자 모델링을 이용하여 분자의 특성을 예측하고, 분자의 특성에 따라 분자를 선별하는 기술로, 분자 모델링 소프트웨어를 사용하여 분자의 구조를 생성하고, 분자의 특성을 예측하는 데 사용된다.

가상 분자 모델링 및 가상 스크리닝은 분자 모델링 소프트웨어를 사용하여 분자의 구조를 생성하고, 분자의 특성을 예측하는 데 사용된다. 분자 모델링 소프트웨어는 분자의 구조를 생성하고, 분자의 특성을 예측하는 데 사용된다. 가상 스크리닝은 분자 모델링을 이용하여 분자의 특성을 예측하고, 분자의 특성에 따라 분자를 선별하는 기술로, 분자 모델링 소프트웨어를 사용하여 분자의 구조를 생성하고, 분자의 특성을 예측하는 데 사용된다.

가상 분자 모델링 및 가상 스크리닝은 분자 모델링 소프트웨어를 사용하여 분자의 구조를 생성하고, 분자의 특성을 예측하는 데 사용된다. 분자 모델링 소프트웨어는 분자의 구조를 생성하고, 분자의 특성을 예측하는 데 사용된다. 가상 스크리닝은 분자 모델링을 이용하여 분자의 특성을 예측하고, 분자의 특성에 따라 분자를 선별하는 기술로, 분자 모델링 소프트웨어를 사용하여 분자의 구조를 생성하고, 분자의 특성을 예측하는 데 사용된다.

- [1] Walters, W. P.; Stahl, M. T.; Murcko, M. A. "Virtual screening" *Drug Discov. Today*, 3(4), 160, (1998).
- [2] Drewry, D. H.; Young, S. S. "Approaches to the design of combinatorial libraries" *Chemmon. Intell. Lab. Sys.* 48, 1, (1999).
- [3] Kubinyi, H. "Similarity and Dissimilarity: A Medicinal Chemist's View" *Perspect. Drug Discov. Design*, 9/10/11, 225, (1998).
- [4] NCI DB, WWW URL <http://chem.sis.nlm.nih.gov/nci3d/>
- [5] MDL mol file, WWW URL <http://www.mdli.com/downloads/literature/ctfile.pdf>
- [6] MMFF94 (Merck Molecular Force Field); Halgren, T. A.; Nachbar, R. B. *J. Comput. Chem.* 17, 587, (1996).
- [7] CFF (Consistent Force Field); Maple, J. R.; Hwang, M. -J.; Stockfisch, T. P.; Dinur, U.; Waldman, M.; Ewig, C. S.; Hagler, A. T. *J. Comput. Chem.* 15, 162, (1994).
- [8] MM2 (Allinger Molecular Mechanics force field version 2); Allinger, N. L. *J. Am. Chem. Soc.* 99, 8127, (1977).
- [9] MM3 (Allinger Molecular Mechanics force field version 3); Allinger, N. L.; Yuh, Y. H.; Li, J.-H. *J. Am. Chem. Soc.*, 111, 8551, (1989).
- [10] Ghang, G.; Guida, W. C.; Still, W. C. "An internal coordinate Monte Carlo Method for searching conformational space" *J. Am. Chem. Soc.*, 27, 1283, (1988).
- [11] Sadowski, J.; Wagener, M.; Gasteiger, J. "CORINA: Automatic Generation of High-Quality 3D-Molecular Models for Application in QSAR" *QSAR and Molecular Modelling: Concepts, Computational Tools and Biological Applications* F. Sanz, J. Giraldo, F. Manaut (Editors), Prous Science Publishers, 646, (1995).
- [12] Bernard, J. M. "Substructure Searching Methods: Old and New" *J. Chem. Inf. Comput. Sci.* 33, 532, (1993).
- [13] Ullman, J. R. "An Algorithm for Subgraph Isomorphism" *J. ACM.* 23, 31, (1976).
- [14] Hall, L. H.; Kier, L. B. "The molecular connectivity chi indexes and kappa shape indexes in structure-property modeling" *Reviews in Computational Chemistry* ed. Lipkowitz, K. B.; Boyd, D. B. Chapter 9, 367, (1992).
- [15] Higo, J.; Go, N. "Algorithm for rapid calculation of excluded volume of large molecules" *J. Comput. Chem.* 10(3), 376, (1989).