한국농약과학회지 제6권 제3호(2002)
The Korean Journal of Pesticide Science
Vol. 6, No. 3. pp.166~174(2002)

〈총 설〉

정량적인 구조-활성상관(QSAR) 기법에 의한 새로운 농약의 개발 I. 기본 개념과 QSAR 기법의 유형

성 낙 도*

충남대학교 농업생명과학대학 응용생물화학부

요약 : 정량적인 분자 구조와 물리-화학적인 성질 사이의 상관관계(QSAR)식을 이용하여 약효성을 예측하고 새로운 농약을 탐색하거나 개발하는데 있어서 효율적인 수단으로 활용되는 QSAR 기법의 발전 과정과 자유 에너지 직선관계(LFER)에 관한 기본 개념을 위시한 QSAR 기법의 목적과 유용성 그리고 장단점과 활용 상 제한점 등에 관한 일반적인 내용에 대하여 간략하게 논의하였다. (2002년 8월 19일 접수, 2002년 9월 30일 수리)

Key words : LFER (Linear free energy relationship), QSARs (Quantitative structure-activity relationships), Agrochemicals development, Basics concepts.

서 론

유기화합물의 구조와 반응성에 관한 자유 에너지 직선관계(LFER)식인 Hammett 식(Johnson, 1975)에 기초하여 정량적인 구조-활성상관(QSAR) 기법의 토대가 되는 Hansch 이론(Hansch 등, 1964)이 발표된 이후, 약 40 여년이 지나는 지금에 이르기까지 기술 선진국에서는 QSAR 기법의 발전을 주도하면서 수많은 의약과 농약들을 개발하여 상용화를 선도하고 있다.

작년도 한국 농약공업협회(KACIA)의 자료에 의하면, 외국산 농약 수입액은 약 3억5천만 달러인 반면에 수출액은 수입액의 약 8.7%인 약 4천만 달러에 지나지 않는 심각한 불균형을 보이고 있으며, 주지하는 바와 같이 외국산 원제만을 복제하여 생산하고 판매에만 급급하던 국내 농약 회사들은 첨단 기술력과 거대한 자본력을 겸비한 외국의 유명 농약 회사들에게 잠식되어가고 있어 독자적인 회생 가능성이 없어 보인다. 이와 같은 종속 관계를 극복하려면, 농약학 분야에 관련한 그간의 획일적인 연구에서 탈피하여 높은 부가가치를 창출할 수 있는 농약 개발에 대한 연구가 더욱 발전되어야 하겠다. 그러므로 새로운 농약의 개발에 관한 연구는 시대적인 요청이기는 하지만 학문적으로는 물론, 제반 기반기술이 구축되어 있지 못한 매우 열악한 현실이다.

활용에 따른 방법론(Blankley, 1983)에 있어서 QSAR 기법은 화합물의 구조 변화에 따른 약리 작용의 변화를 분자설계 시스템을 이용하여 정량적으로 분석 할 뿐만 아니라, 약리 작용을 분자 구조와 물리-화학적인 성질과의 상관관계(Chapman 등, 1978)식을 유도하여 약효가 가장 큰 물질을 예측하고 합성함으로써 합리적이고도 효율적으로 새로운 농약을 개발하는 방법이며(Baker 등, 2001), 정밀화학 분야에 있어서 다양하게 이용되는 강력한 도구이자(Reddy 등, 1999) 수단이다. 왜냐하면, 구조-활성 상관관계에는 생물활성을 발현하는 화합물의 구조적인 요소들과 생물활성과의 관계에 있어서 생물활성의 특징을 작용기구(mode of action)의 특징으로 규정할 수 있기 때문이다.

저자는 살충성 유기인계 화합물에 따른 acetylcholinesterase의 인산화(성, 1984 및 1986)와 carbamate계 화합물의 carbamylation(성, 1989)에 대한 LFER과 QSAR 기법에 관한 논리를 국내 농약학계에 처음으로 소개한 이래로, 새로운 농약의 탐색 연구(Sung, 2000)를 지속하고 있다. 때늦은 감이 있으나 새로운 농약의 탐색과 개발에 관련한 QSAR 기법의 이해와 활용을 권장하기 위하여 본 총설, (I)에서는 QSAR 기법(ACCVIP)의 기본 개념과 목적, 활용상의 장·단점 그리고 유형에 따른 일반적인 내용에 대하여 간략하게 논의하고 소개하고자 한다.

*연락저자

QSAR의 역사와 기본개념

지금으로부터 134년 전인 1868년 Crum-Brown과 Fraser에 의하여 alkaloid들의 생리활성에 관한 초보적인 QSAR 연구가 시작된 이래로(표 1) 1899년 Meyer와 Overton은 마취작용에 관한 lipoid theory를 발표하여 생리활성 크기에 지질과 물 분자 사이의 분배계수인 소수성(hydrophobicity)이 중요한 의미를 가진다는 사실을 처음으로(Tute 1990) 확인하였다.

1930년대에 Hammett는 벤조산(R-H) 유도체들의 치환기(R-X) 변화에 따르는 산-해리 상수에 대한 비의 대수를 sigma 상수로 정의함으로서 반응에 기초한 자유 에너지 직선관계(LFER)로부터 전자적인 성질이 평형상수(K)는 물론, 산-해리 상수(pKa)와 정량적인 관계가 있다는 아래와 같은 Hammett 방정식(Johnson, 1980)을 제안하였다. 여기서 sigma(σ) 상수는 방향족 화합물에 국한하여 사용할 수 있는 치환기 상수로서 비 치환-phenyl 고리상 H 원자를 기준으로 치환기의 전자 수수정도에 관한 상수이고 rho(ρ)는 반응의 감도를 나타내는 비례 상수이다.

$$\log K_{R\text{-}X}-\log K_{R\text{-}H}=\rho\sigma$$

이러한 LFER식에 기초하여 1962년에 이르러 Hansch는 친 지질성인 소수성을 포함시켜 화합물의 생물학적인 유용성을 설명(Hansch, 1969)하였다. 따라서 약물과 수용체간의 상호작용에 의한 생물활성(log1/C)은 다음 식과 같이 크게 전자 전달효과(electronic effect; σ)와 입체효과(steric effect; Es) 그리고 소수성 효과(hydrophobic effect; π) 등의 3가지 형태로 분류되는 물리-화학적인 성질(설명인자)들에 대한 선형 대수의 합으로 설명함으로서 약제 설계에 활용할 수 있는 QSAR기법에 대하여 초열역학적 접근(extrathermodynamic approach)이 이루어 졌다.

$$\log 1/C=a\pi^2+b\pi+\rho\sigma+cEs+d$$

이와 같이 QSAR식은 보통 다중 선형 회귀식(MLR: multiple linear regression)과 같은 통계적인 방법을 사용하여 물리-화학적 성질은 독립 변수로 그리고 생물활성(BA; biological activity)들은 종속 변수로 아래의 식과 같이 표현됨으로서 본격적으로 2D QSAR에 대한 연구가 시작되는 토대를 구축하게 되었다. 이 식에서 BA항은 실험적으로 측정되는 기질 물질의 농도(log1/C), pI_{50}, IC_{50}, ED_{50} 및 Km 등과 같은 척도인 반면에 우변 항의 독립 변수들은 수용체와의 상호작용에 관여하는 기질 분자의 위상학적인 상수나 물리-화학적인 성질을 나타내는 진술한 바와 같은 3 가지 형태의 설명인자(descriptor) 들이다.

$$BA=a_1x_1+a_2x_2+ \ldots\ldots\ldots +a_nx_n$$

1988년에 Cramer 등은 Hansch 식에서 언급된 3가지 물리-화학적인 성질을 비교 분자장으로 표현하는 3D QSAR인 CoMFA 기법(Cramer 등, 1988)을 제안하여 현재 가장 보편적으로 활용되는 방법으로 인식되기에 이르렀다. 1997년에는 2D 화합물의 구조적 특징들과 생물활성 사이의 예측관계를 부분 최소자승(PLS; patial least square) 방법으로 분석하는 분자 홀로그램(H) QSAR기법(Heritage 등, 1999) 그리고 Hopfinger 등에 의하여 3D QSAR 기법의 2가지 한계점 즉, 반응점의 가정과 유연한 생물 조절제를 표현하는데 있어서 평준화된 수용체 모델의 유용성 등을 보강한 리간드 분자의 양성자화 상태, 다중 형태 및 지향성이 고려된 개념의 4D QSAR 기법(Hopfinger 등, 1977)이 알려졌다. 또한, 최근에 Vedani 등은 induced-fit 가설(5차원)의 다중적인 설명뿐만 아니라, 리간드의 위상학적인 복식표현 개념의 5D QSAR기법(Vedani 등, 2002) 까지 등장하기에 이르렀다.

모든 QSAR식은 방법과 표현은 다르다 할지라도 구조상의 요소와 생물활성 크기에 대한 2가지 변화들 간의 대응 논리를 근간으로 하고 있으며 그림 1에는 QSAR 개념의 과학적인 측면들에 대하여 요약하였다. 따라서 QSAR기법은 활성 메카니즘을 규명하려는 생물학적인 실험과 기질분자-수용체 사이의 구조와 상호작용에 관한 화학적인 이론 및 이들 두 요소들의 분석방법으로서 통계학적인 경험 등, 3가지 유형의 측면들을 망라하는 최적화 수단으로서 QSAR의 기본개념은 다음 식과 같이 표현된다. 여기서 x_i는 구조적인 성질들로 최소자승 회귀분석으로 얻어지는 값이고 a_i는 x_i의 계수이며 m은 화합물의 수이다.

$$BA=\sum_{i=1}^{m}a_ix_i+Const$$

모든 화학반응의 속도 상수와 온도와의 관계는 Arrhenius 식(kr=A exp(-E≠/RT))으로 잘 정리되어 있으며 Hammett 관계식들은 모두 1차식(logk=alogK+b)의 관계를 나타낸다. 따라서 온도가 일정한 조건에서 이들 두 관계식을 짝지워 정리하면 다음 식과 같이 자유 에너지 변화인 $\triangle F^{\neq}$와 $\triangle F^{\circ}$사이에는 직선관계가 성립한다. 그러므로 Hammett식이 성립할 때, 자유에너지 직선관계(LFER; $\triangle G=-RT\ln K$)가 존재한다

Table 1. History & Development of QSAR methodologies.

Year	Authors	Comments
1868	Crum-Brown & Fraser	First general formulation of QSAR.
1869	Richardson	Narcotic effect of primary alcohols.
1899~1901	Meyer & Overton	Formulated the lipoid theory of narcosis.
1939	Ferguson	Formulated partition coefficient and thermodynamics.
1932	Hammett	Hammett equation proposed.
1956	Bruice, Kharasch & Winzler	QSAR Methodology rapidly developed.
1962~1964	Hansch & Fujita	Birth of modern 2D-QSAR Mothodology.
1988	Cramer	3D-QSAR (CoMFA) was formulated.
1997	Hopfinger et. al.	4D QSAR, Ligand flexiblity
2001	Vedani et. al.	5D-QSAR., Induced-fit. Receptor flexibility

고 하며 이 식의 성립조건은 한 계열의 반응 메카니즘이 같고 구조가 대단히 유사한 경우라고 말할 수 있다. 따라서 LFER은 유사성의 정량적인 모델이라고 번역될 수 있다.

$$\triangle F^{\neq} = a \triangle F^{\circ} + \text{Constant}$$

이상과 같이, 화합물이 생체내에 흡수된 후의 활성 발현은 작용점에 이르는 수송과정과 작용점에서 일어나는 상호작용 그리고 그 결과로 인하여 관측되는 생물활성을 발현하는 과정 등, 3가지 과정으로 구분되며 QSAR 기법은 구조상 요소의 변화와 활성 변화에 대응하여 발전되고 있다.

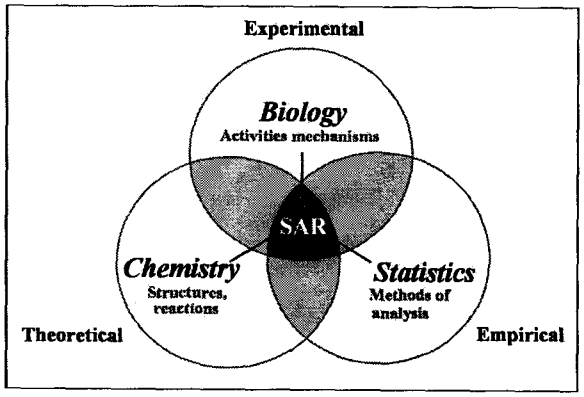

Fig. 1. Applicable R&D criteria of quantitative structure-activity relationships (TOPKAT).

왜, QSAR 기법이 필요한가?

신 농약 개발의 모태로서 약 화학은 생명 관련 산업의 중추적인 학문 분야이며 물리-유기화학을 기본 바탕으로 하는 약제 개발수단(Hansch, 1995)으로서 QSAR 기법을 이해하고자 함은 약리 효과가 가장 큰 화합물을 탐색하는 데에 따른 비용과 시간 및 노력 등을 최소화함으로써 약물개발 연구에 있어서 가장 효율적인 방법이기 때문이다.

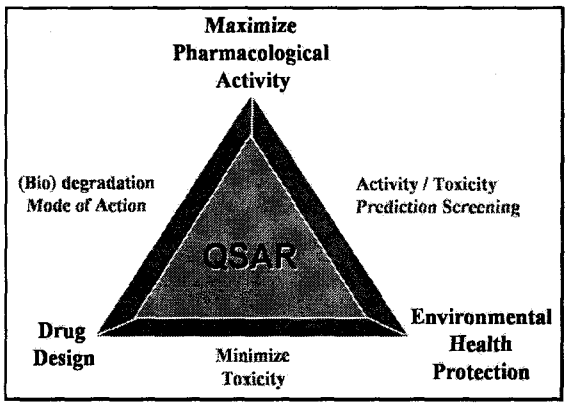

Fig.2. Relationships of three factors, activity dru design, and environmental and heal protection in QSAR (Waterbeemed, 1996).

먼저, QSAR 기법의 구체적인 연구 목적(Tute 1990)은, 그림 2에 제시된 바와 같은 계통도에 의하여 즉, 약효검정 대상 기질 화합물 군들이나 분자 내 특정 치환기의 물리-화학적 성질과 활성의 형태 및 생물활성 사이의 상관성을 발전시킴으로서 구조상 요소의 변화와 활성 변화와의 두 가지 측면인 생체계의 다양성과 불균일성 그리고 생리활성 물질의 구조적 특이성에 따른 화학반응이나 물리-화학적 평형을 지배한다는 원칙에 기초한 생리활성 요소나 활성 발현 메카니즘(또는 생 분해나 작용기작)에 대한 단서를 이해하

는 데에 있다. 그러므로 약리학적 활성은 극대화하되 환경 독성이나 잔류성을 최소화하거나 그리고 높은 활성과 선택성 및 약효 지속성을 조절하고 다루지 않은 화합물에 대한 활성을 예측하며 약물 분자설계 (Ramsden, 1990) 구축을 용이하게 할 뿐만 아니라, QSAR 기법을 활용한 연구 결과들로부터 생물환성이 가장 우수한 물질을 도출하는 일련의 작업을 반복함으로써 유효한 화합물(hit compound)을 얻기 위함이다.

다음은 QSAR 기법의 3가지 특징적인 유용성과 다양한 필요성(표 2)에 대하여 요약하였다. 첫째로는 생물활성 실험의 질적인 신뢰성과 자료에 대한 합리성을 추구하며, 둘째로는 실험 결과의 모순점은 물론, 실험값과 모델에 의하여 예측된 값들 사이의 편차를 알 수 있고 투약량 범위를 평가하여 진단하며 그리고 셋째로는 분자 설계 시 독성은 최소화하되 활성은 극대화할 수 있을 뿐만 아니라, 아울러 그 성질들을 예측하는 일이다. 그리고 QSAR 기법을 수용하여야 하는 이유로서, QSAR 기법의 필요성은 다음과 같다. 예컨대 수많은 자료를 분석하고 검토하는 데에 필수적인 컴퓨터 기술이 발달하였고, 많은 화합물의 합성과 그에 따른 생물활성을 검정하는데 있어서 실험량을 줄여야 할 필요성이 대두되었으며, 생물활성 메카니즘과 수용체의 구조와 성질 및 대사 등에 관련한 지식수준의 량과 질이 크게 향상됨에 따른 학문적인 체계를 확립하고 합리성을 추구하여야 할 필요성이 증대되었기 때문이다. 뿐만 아니라, 많은 개발비용과 약리 효과와 효율성이 결여된 약제들로 인한 약해 등, 경제적인 피해에 대한 사회적인 압력을 완화하거나 축소시켜야 하는 당위성과 보다 향상된 인체의 건강과 자연환경을 유지 내지는 개선하기 위한 고도의 선택성과 환경 친화적인 약제 개발의 필요성을 충족하기 위하여 QSAR 기법의 활용은 불가결한 선택일 수밖에 없다. 이러한 당위성에 관련하여 그림 3에는 QSAR 기법의 최적화 과정중 약제의 설계, 생물 활성시험 및 합성 등, 3가지 요소 사이에 내재되어 있는 논리상 개념의 관계를(Waterbeemd, 1996) 제시하였다.

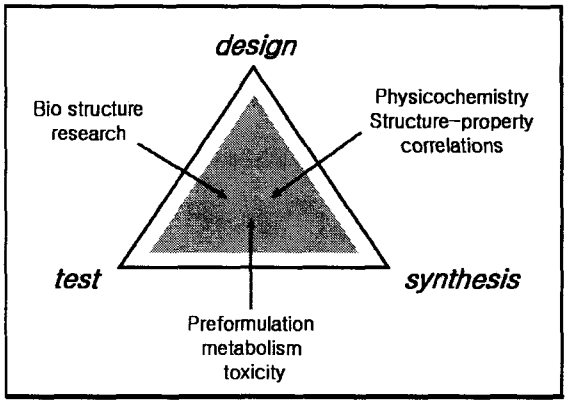

Fig.3. The lead optimization processof QSAR The classical approach involved mainly th cycle design, synthesis and testing.

이상과 같이 구체적이고 체계적일 뿐 아니라, 신속하고 정확하게 기질과 수용체 사이의 관계를 이해할 수 있도록 함으로서 새로운 농약을 개발하는데 있어서 QSAR 기법의 도입(Draber 등, 1992)과 활용(Ramsden 1990)은 선택이 아니라, 필수일 수밖에 없다.

Table 2. Use and need in utilizing QSAR methodology

Use	Necessity
Rationalize existing data.	Advances in computer technologies.
-quality assurance of experiment.	Reduce need for synthesis & biological testing.
-provide insight into underlying mechanism of action.	Increase need for rationalizing the advances in knowledge.
-generate mechanism hypotheses.	-molecular mechanisms of biological activities.
Direct experiment.	-receptors, metabolism, etc.
-dose range estimates.	Economic and social pressures.
-suggest chemicals testing.	-reduce costs.
-point to inconsistencies in experimental results (i.e. errors).	-reduce toxicity of residues in crops.
Predict.	Better health and environmental protection.
-hazard identification screening.	-strategic application of limited testing resources.
-minimize toxicity in chemical design.	-screen larger numbers of chemicals more efficiently.

Table 3. Classification in application fields and limiting factors of QSAR

Type	Application fields	Limiting factors
QSAR	Syntheses of fine chemicals.	Data availability.
	Searching and development of drug & biological active materials etc.	Knowledge of mechanisms of activity.
	Reaction mechanism & degradation.	Chemical classification.
		-Molecule might contribute to the potency.
QSTR	Environmental toxicology/residue	Molecule representation.
QSBR	Mode of action and biodegradation.	Model interpretation & validation
QSPR	Various physico-chemical properties & constants.	-Colinearity among different parameters.

응용분야 및 장점과 단점

QSAR 기법의 유형은, 기질 화합물의 구조를 평면(2D) 구조로 간주하는지 아니면 입체적인 3차원(3D) 구조를 대상으로 연구하는가에 따라 2D QSAR과 3D QSAR로 각각 구분된다. 그러나 4D 및 5D QSAR 기법은 이들보다 하위 단계 기법의 제한점을 개선하여 질적인 수준 향상에 따른 구분일 따름이다.

QSAR 기법의 기본인 2D QSAR 기법은 화합물의 특정한 생물 활성값과 물리-화학적인 성질로 구성된 자료를 만들어 다중 선형회귀(MLR; multiple linear regression) 분석법으로 모델을 구축하여 약리작용 결과를 해석하거나 확인하고 예측한다. 그러나 3D QSAR은 화합물의 안정한 형태를 탐색하여 정하고 3차원 공간상에 정렬하여 설명 인자들을 계산하고 보통 부분 최소자승법(PLS)으로 통계 처리하여 모델을 구축한다. 2D QSAR 기법 중에는 Classical QSAR이라고 불리우는 Hansch 식에 기초한 2D QSAR(Kubinyi, 1993a) 및 분자 홀로그램(H) QSAR(Heritage 등, 1999) 그리고 3D QSAR 기법(Kubinyi, 1993b) 중에는 CoMFA(Cramer 등, 1988) 및 CoMSIA(Klebe 등, 1994) 등의 기법들이 근래 가장 보편적으로 대중화되어 있는 방법이다.

정량적인 구조와 반응성 관계는 다루어지는 분야에 따라(표 3), 구조와 생물 활성과의 관계를 다루는 QSAR 기법(Kubinyi, 1993a)을 위시하여 구조와 물리-화학적인 성질과의 관계는 QSPR(Kansy, M. (1996), 구조와 독성과의 관계는 QSTR(TOPKAT) 및 구조와 생분해 반응과의 관계를 다루는 분야는 QSBR 등으로 각각 구분한다.

한편, QSAR기법을 적용하는데 있어서 제한점(표 3)을 살펴보면, 약제의 생체내 흡수 및 이행에 관한 분배속도 및 분배율, 반응 점에서 약제와 수용체간의 입체적인 상호작용 및 친화력에 대한 탈 용매화 에너지의 영향, 수소결합의 세기 등을 포함하여 보다 구체적인 약리작용 등, 자료의 유용성, 활성 기여정도에 따른 화합물의 분류, 분자에 대한 표현, 모델의 해석 및 타당성 여부를 이해하기가 용이하지 않다는 점들이 있다. 그러나 이러한 문제점들은 보다 구체적인 설명인자들을 연구하고 탐색함으로서 해소될 수 있을 것이다. 또한, 2D 및 3D QSAR 기법들에는 다음과 같은 장, 단점(Kim, 1993)들이 있으므로 잘 파악하여 서로 상호 보완적으로 사용하면 보다 정확하게 새로운 물질을 예측하는데 큰 도움을 받을 수 있다.

2D QSAR 기법의 장점(표 4)은, 운영하기가 간편하고 계산 시간이 오래 걸리지 않기 때문에 값 비싼 컴퓨터 시설이 요구되지 않으며 여러 물리-화학 파라미터로서 설명 인자인 독립 변수들을 그 관계식에 적절히 반영하여 생물 활성의 세기와 여러 변수간의 상관성(r^2)이 높은 최적의 QSAR식을 얻을 수 있다. 다양하고 구체적인 설명인자들을 활용하여 생물활성을 심도있게 이해하고 활성의 수준을 평가할 수 있을 뿐만 아니라, 자료의 최적화를 통하여 합성 화합물의 수를 현저히 줄일 수 있으며 활성에 미치는 물리-화학적인 요인이나 생물활성 등을 외삽법(extrapolate) 또는 내삽법(interpolate)으로 예측하거나 기질과 수용체간의 상호작용과 작용 메카니즘을 진단하는 수단 등으로 활용된다. 또한, 3D QSAR 기법에서 가장 중요시되는 안정한 형태의 공간상 정렬 같은 절차가 요구되지 않으며 모든 결과들이 방정식으로 간편하게 정리된다.
반면에 단점은, 일반적으로 계산이 복잡하며 생물활성(MIC와 같은 측정치)의 신뢰도(또는 정확도)가 좋지 않을 경우에는 실제와 다른 의외의 엉뚱한 결과가 나올 가능성이 크며, 동일한 유도체들이 아닌 서로 상이한 구조를 가지는 화합물들은 적용이 불가능 할 뿐만 아니라, chance correlation에 따른 우를 범할 가능성이 있으며 실제 3차원 구조에 대한 개념이 포함

되지 않아 입체 효과를 효율적으로 이해할 수 없고 더욱이 수소-결합 같은 부분에 대한 정보를 얻을 수 없다는 점 등이다.

3D QSAR 기법(Kubinyi, 1993b)의 장점으로는 2D QSAR 경우와 달리, 구조가 약간 상이한 일련의 화합물의 경우에도 적용이 가능하여 입체적인 구조로 입체장, 정전기장 및 소수성장 등의 결과들이 등고도(Contour map)상에 computer graphic으로 표현되므로 입체효과와 수소결합에 의한 효과를 2D QSAR의 경우보다 잘 확인할 수 있다. 또한, 다양한 설명 인자가 요구되지 않으며 chance correlation을 걱정 할 필요가 없다. 반면에 단점으로는, 값비싼 컴퓨터 시설이 요구되며 3차원적인 형태와 겹침 법칙에 따라 3차원 공간상에 분자 구조들을 정렬 시켜야만 한다. 다루지 않은 부분에 대한 외삽(extrapolate)에 의한 예측이 불가능하며 상관식의 변수들로부터 얻을 수 있는 정보가 2D QSAR 기법보다 적다는 단점이 있다. 따라서 2D-QSAR 및 3D-QSAR 기법들 사이는 상호 보완적이며 서로 다른 장단점을 가지고 있을 따름이다.

2D QSAR과 3D QSAR

새로운 농약을 개발하는데 있어서 보편화되어 있는 2D QSAR과 3D QSAR 기법에 관한(Blankley, 1983) 실행(Accelrys) 과정에 대하여 설명하고자 한다.

일반적으로 2D QSAR 기법은 먼저, 대표적인 치환기가 치환된 한 조의 화합물들을 합성하여 생물활성을 측정한 다음, 각 화합물에 대하여 생물활성과 치환기들에 관한 설명 인자들과의 관계를 회귀분석(regression analysis) 방법(Dillon 등, 1984)으로 계산하여 합성한 전체 화합물에 대하여 상관성이 높은 QSAR식을 얻는다. 이 식으로부터 여러 설명인자인 독립 변수들의 종류와 계수(weighting factor)의 크기를 감안하여 치환기와 생리활성 사이의 상관관계에 미치는 경향을 파악한다. 그 경향에 따라서 생물활성 향상에 기여 할 것으로 유력시되는 치환기들을 추가로 선택하여 다시 일련의 화합물들을 합성하고 활성을 측정한 후에 처음 시도하여 얻은 QSAR식과 함께 연립 방정식을 세우고 회귀분석을 통하여 상관성 큰, 예측성이 향상된 QSAR식을 유도(Hansch 등, 1995)한다. 이 식으로부터 생물활성이 가장 좋을 것으로 예상되는 화합물을 선택하여 합성하는 이러한 일련의 방법을 반복하여 생물활성이 높은 목적하는 화합물을 얻는다.

이상과 같은 2D-QSAR 결과와 방법을 보완하는 차원에서 3D QSAR 기법(Kubinyi, 1993b)은 대략 다음과 같은 과정으로 실행된다. 먼저, 기질 분자의 생물활성과 물리-화학적인 성질에 관한 자료를 만들고 안정한 형태를 결정하여 3차원 공간상에 겹침 법칙에 따라 정렬(geometry optimization)한다. 설명인자를 계산하고 통계 분석하여 QSAR 모델을 구축한 다음에 결과를 해석하고 확인하며 예측하는 일련의 과정을 거친다. 그러므로 3D QSAR은 2D QSAR 기법에서 가장 취약점으로 인정된 구조적인 해석상의 문제점을 크게 보완하였다는 점이다.

이 같은 맥락에 따라. 3차원의 구조 개념이 보강된 3D-QSAR 기법 중에서 가장 유용하게 활용되고 있는 비교 분자장 분석(CoMFA) 기법(Cramer 등, 1988)이 제안되었다. CoMFA는 3차원 공간에 일정한 간격으로 놓여있는 격자 원자의 입체장(steric field)과 정 전기장(electrostatic field) 그리고 소수성 결합 성분을 고려한 HINT(hydropatic interaction techique) (Goodford, 1985)를 도입하여 분자의 친 지질장(lipo-

Table 4. Advantages and disadvantages of 2D QSAR methodology (Kim, 1993)

Advantages	Disadvantages
Simple to run.	Usually applicable only to a congeneric series.
Can be applied for drug-receptor interaction.	Parameters may by missing or estimated.
Require little computing time.	Usually dose not include 3D structural information.
Does not require an expensive computer facilities.	
Does not require the bioactive conformation.	Needs to watch for chance correlation.
Does not require a superposition rule.	Dose not describe steric effects well.
May extrapolate into unexplored region with care.	Dose not describe hydrogen-bonding effects well.
Results are summarized in an equation.	
No adjustable parameters are involved.	Needs to watch for chance correlation.

philic potential field)인 소수성 장(hydrophobic field) 등, 3가지 비교 분자장으로 표현되는 등고도(contour map)로부터 분자의 물리-화학적인 성질과 생물활성 사이의 상관성을 연구하는 기법이다.

설명인자가 많은 CoMFA의 경우에는 주성분 회귀분석(PCR; principal component regression analysis)과 부분 최소자승(PLS) 법(Stahle 등, 1988)으로 활성과 설명 인자간의 상관 관계식을 유도하는데 사용되며 회귀 모델로서 적합성 여부를 판단하고 모델의 예측성을 평가한다. 따라서 3D-QSAR 기법이 2D-QSAR 기법보다 비교 우위적으로 양호하다는 개념이 아니며 이들 QSAR에 관련한 전문 학술지(QSAR)가 격월간으로 발간되고 있다.

다음 총설(II)에서는 QSAR기법을 구체적으로 이해하고 활용하는데 있어서 필요한 LFER과 설명인자들의 분류에 따른 연구사례는 물론, 신 농약 개발과정 등에 대하여 알아보고자 한다.

인용문헌

Accelrys: Cerius^2 QSAR+, Insight II, Catalyst program package (Imagene). Molecular Simulations Inc. & www.sinica.edu.tw

ACCVIP; The Australian Computational Chemistry via the Internet Project, QSAR, www.chem.swin.edu.au

Baker, D. R. and N. K. Umetsu (2001) Modern Agrochemical Discovery. *In* Agrochemical Discovery; Inset, Weed, and Fungal Control (ed. Baker, D. R. and N. K. Umetsu). Ch.1., ACS Symposium Series 774, American Chemical Society, Washington, DC.

Blankley, J. (1983) Introduction: A Review of QSAR Methodology. *In* Quantitative Structure-Activity Relationships of Drugs (ed. Topliss, J. G.). Academic Press, New York.

Chapman, N. B. and J. Shorter (1978) Correlation Analysis in Chemistry: Recent Advances. Plenum Press, New York and London.

Cramer, III, R. D., D. E. patterson and J. D. Bunce (1988) Comparative molecular field analysis (CoMFA), I. Effect of shape on binding of steroids to carrier proteins, J. Am. Chem. Soc. 110:5959~5967.

Dillon, W. R. and M. Goldsten (1984) Multivariate Analysis: Methods and Applications. pp.209., John Wiley & Sons, New York.

Draber, W. and T. Fujita (1992) Rational Approaches to Structure, Activity, and Ecotoxicology of Agrochemicals. CRC Press, Boca Raton.

Goodford, P. J. (1985) A computational procedure for determining energetically favorable binding site on biologically important macromolecules, J. Med. Chem. 28:849~857.

Hansch, C. and T. Fujita (1964) Rho-sigma-pi analysis. A method for the correlation of biological activity and chemical structure, J. Am. Chem. Soc. 1616~1626.

Hansch, C. (1969) Quantitative approach to biochemical structrure-activity relationships, Acc. Chem. Res. 2:232~239.

Hansch, C. and A. Leo (1995) Exploring QSAR: Fundamentals and Applications in Chemistry and Biology. ACS Professional Reference Book, American Chemistry Society, Washington, DC.

Heritage T. W. and D. R. Lowis (1999) Molecular Hologram QSAR. Ch. 4., *In* Rational Drug Design :Novel Methodology and Practical Applications (ed. Parrill, A. L. and M. R. Reddy), ACS Symposium Series 719, American Chemical Society. Washington, DC.

Hopfinger, A. J., S. Wang, J. S. Tokarski, B. Jin, M. Albuquerque, P. J. Madhav and C. Duraiswami (1997) Construction of 3D QSAR model using the 4D QSAR analysis formalism. J. Am. Chem. Soc., 119:10509~10524.

Johnson, C. D. (1975) Linear Free Energy Relationships and Reactivity-Selectivity Principle, Chem. Rev. 75:755~765.

Johnson, C. D. (1980) The Hammett Equation. Cambridge University Press, Cambridge.

KACIA; Korea Agrochemical Industry Association., http://www.kacia.or.kr

Kansy, M. (1996) Molecular Properties, Ch.2., *In*

Structure-Property Correlations in Drug Research (ed. Waterbeemd, H. V. D.). Academic Press, R. G. Landes Co. Austin.

Kim, K. H. (1993) Comparison of Classical and 3D QSAR. pp.619~642, *In* 3DQSAR Drug Design Theory, Methods and Applications (ed. Kubinyi, H.) ESCOM. Leiden.

Klebe, G., U. Abraham and T. Mietzner (1994) Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their bioligical activity, J. Med. Chem. 37:4130~4146.

Kubinyi, H. (1993a) QSAR, Hansch analysis and related approaches, *In* Methods and Principles in Medicinal Chemistry (ed. Mannhold, R., L. P. Krogsgaard and H. Timmerman), Vol. 1., VCH, Weinheim.

Kubinyi, H. (1993b) 3D QSAR in Drug Design; Theory, Methods and Applications. ESCOM, Leiden.

QSAR: Wiley-VCH Verlag GmbH, P. O. Box 101161., D-69451 Weinheim, Germany.

Ramsden, C. A. (1990) Quantitative Drug Design. *In* Comprehensive Medical Chemistry; The Rational Design, Mechanistic Study & Therapeutic Applications of Chemical Compounds (ed. Hansch, C., P. G. Sammes and J. B. Taylor) Vol.4., Pergamon Press, Toronto

Reddy, M. R. and Parrill, A. L. (1999) Overview of Rational Drug Design. *In* Rational Drug Design (ed. Reddy, M. R. and Parrill). Ch. 1., ACS Symposium Series 719, American Chemical Society, Washington, DC.

Stahle, L. and S. Wold (1988) Multivariate data analysis and experimental design in biomedical reearch. Progr. Med. Chem., 25:292~334.

Sung, N. D. (2000) Collected Works. Reaction Kinetics & Mechanism; Exploring QSARs. pp.305~640, OB Planning Co.

TAPKAT; Toxicity prediction by komputer assisted technology program (Ver. 5.0)., Health Designs Inc., 183 E. Main St., Rochester, New York, 14604, USA.

Tute, M. S. (1990) History and Objectives of Quantitative Drug Design, 17.1., *In* Comprehensive Medicinal Chemistry (ed. Hansch, C., P. G. Sammes and J. B. Taylor). Pergamon Press, Toronto.

Vedani, A. and M. Dobler (2002) 5D-QSAR: The key for simulating induced fit? J. Med. Chem., 45:2139~2149.

Waterbeemd, H. V. D. (1996) Design of Bioactive Compounds, Ch.1., *In* Structure-Property Correlations in Drug Research (ed. Waterbeemd, H. V. D.). Academic Press, R. G. Landes Co. Austin.

성낙도 (1984) 살충성 O,O-diethylphenylphophate 유도체들에 의한 Acetylcholinesterase의 phosphorylation에 미치는 자유에너지 관계. 충남대 농업과학연구지 11:176~181.

성낙도, 명평근, 유병태, 이천배 (1986) 살충성 O,O-diethylphenylphophate 유도체들에 의한 Acetylcholinesterase의 phosphorylation에 관한 양자 약리학적 연구. 충남대 약학연구지 2:7~15.

성낙도 (1989) 왜, m-methyl 치환 살충제들은 강한 살충작용을 나타낼까? Phenyl N-methylcarbamate 와 m-xylyl-N-methylcarbamate 유도체들에 관하여. 한국농화학회지 32:170~177.

Development of new agrochemicals by quantitative structure-activity relationship (QSAR) methodologies. I. The basic concepts and types of QSAR methodologies.

Nack-Do Sung*(*Dividion of Applied Biology & Chemistry, College of Agriculture & Life Sciences, Chungnam National University. Daejon 305-764, Korea.*)

Abstract : The fundamental concepts on the basis of linear free energy relatioship (LFER), history of development, prediction of pharmacological effects, advantages and disadvantages, etc. according to the 2D and 3D QSAR methodologies were summarized in utilizing the quantitative structure-activity relationship (QSAR) techniques for searching and development of new agrochemicals. Objectives, role of QSAR techniques in development process of pesticides and limitations in QSARs were discussed and introduced.

*Corresponding author (Fax : +82-42-825-3306, E-mail : ndsung@cnu.ac.kr)