

# QSAR

# QSAR

- Correlazione tra attività biologica di una serie di composti e i loro parametri chimico-fisici in maniera quantitativa utilizzando formule matematiche
- **Misure quantitative** di proprietà fisico-chimiche e biologiche
- **Proprietà fisico-chimiche**

- Idrofobicità della molecola
- Idrofobicità di sostituenti
- Proprietà elettroniche dei sostituenti
- Proprietà steriche di sostituenti

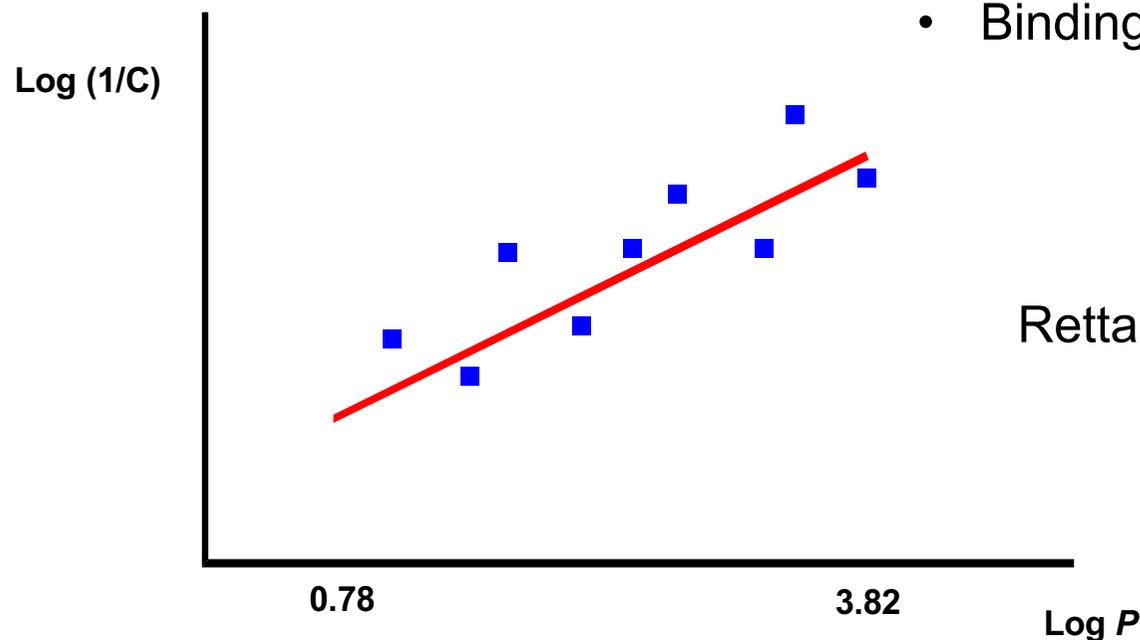
} Proprietà più comunemente studiate

# Idrofobicità di una molecola

$$\text{Coefficiente di partizione } P = \frac{[\text{Farmaco}]_{\text{ottanolo}}}{[\text{Farmaco}]_{\text{acqua}}}$$

*P elevato*  
  
 Elevata idrofobicità

**Attività di farmaci spesso legata a *P*** (e.g. legame del farmaco all'albumina)



- Aumento del binding con l'aumento del log P
- Binding maggiore per farmaci idrofobici

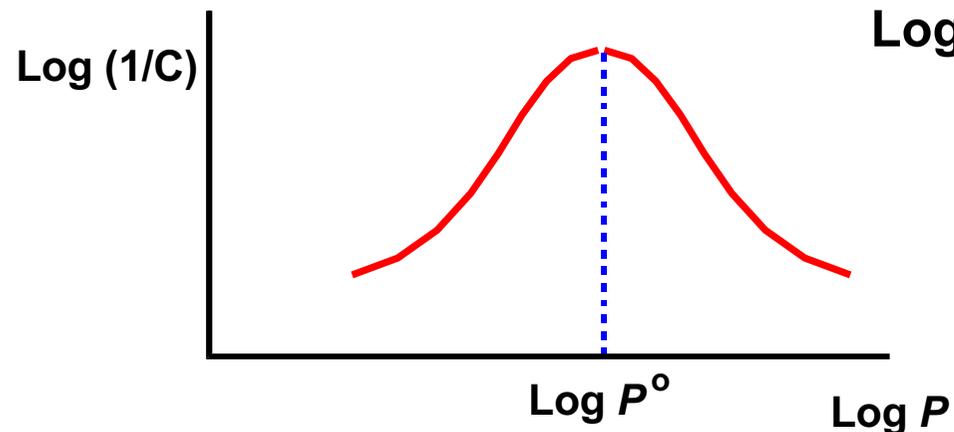
$$\text{Log} \left( \frac{1}{C} \right) = 0.75 \log P + 2.30$$

Retta per un limitato intervallo di log *P*

# Idrofobicità di una molecola

## Attività anestetica generale di eteri

(curva parabolica - più ampia gamma di valori di log P)



$$\text{Log } (1/C) = -0.22 (\text{log } P)^2 + 1.04 \text{ log } P + 2.16$$

**Valore ottimale di log  $P$  per l'attività anestetica = log  $P^{\circ}$**

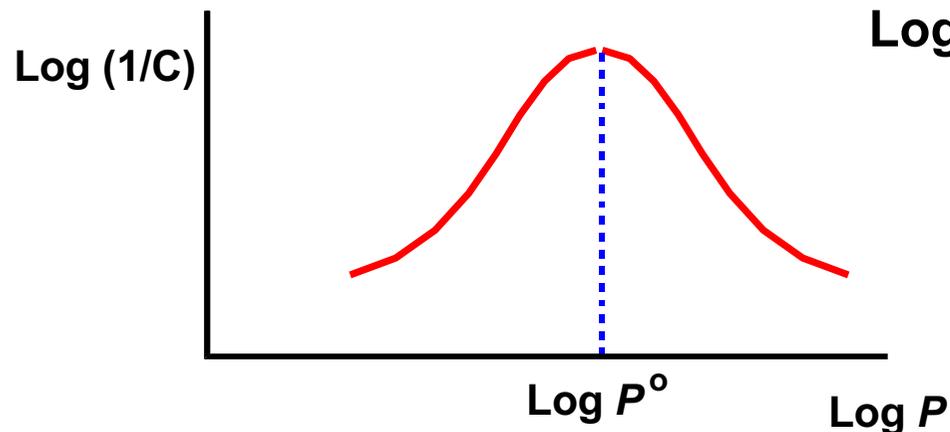
Quando  $P$  è piccolo, il termine  $(\text{log } P)^2$  è molto piccolo e l'equazione è dominata dal termine  $\text{log } P$ . Questa rappresenta la prima parte del grafico in cui l'attività aumenta con l'aumentare  $P$ .

Quando  $P$  è elevato, il termine  $(\text{log } P)^2$  è più significativo e “annulla” il termine  $\text{log } P$ . Ciò rappresenta l'ultima parte del grafico in cui l'attività scende all'aumentare  $P$ .

# Idrofobicità di una molecola

## Attività anestetica generale di eteri

(curva parabolica - più ampia gamma di valori di log P)



$$\text{Log } (1/C) = -0.22 (\text{log } P)^2 + 1.04 \text{ log } P + 2.16$$

Valore ottimale di log  $P$  per  
l'attività anestetica = log  $P^0$

**Equazioni QSAR: applicabili solo a composti della stessa classe strutturale**  
(i.e. eteri).

Log  $P^0$ : simile per anestetici di diverse classi strutturali (circa 2.3)

Strutture con logP ca. 2.3: capacità di accedere al sistema nervoso centrale facilmente (i.e. barbiturici potenti logP di circa 2.0)

Modifiche del valore di logP di farmaci per evitare gli effetti collaterali a carico del sistema nervoso centrale (logP molto diverso da 2.0)

# Idrofobicità dei sostituenti

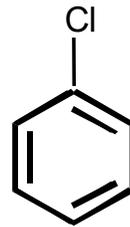
## CONSTANTE d'IDROFOBICITÀ dei SOSTITUENTI ( $\pi$ )

- Misura dell'idrofobicità relativa di un sostituyente rispetto all'idrogeno
- Valori tabulati per sostituenti alifatici e aromatici
- Misurato sperimentalmente con confronto dei valori di logP con il logP della struttura parente

$$\pi_X = \log P_X - \log P_H$$

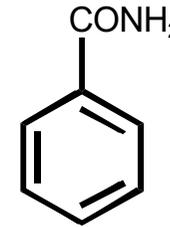


Benzene  
(Log P = 2.13)



Chlorobenzene  
(Log P = 2.84)

$$\pi_{\text{Cl}} = 0.71$$



Benzamide  
(Log P = 0.64)

$$\pi_{\text{CONH}_2} = -1.49$$

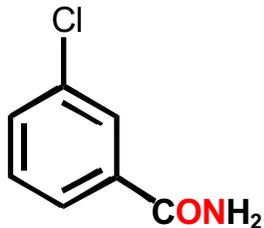
- Valori positivi: sostituenti più idrofobici di H
- Valori negativi: sostituenti meno idrofobici di H

# Idrofobicità dei sostituenti

## CONSTANTE d'IDROFOBICITÀ dei SOSTITUENTI ( $\pi$ )

	CH <sub>3</sub>	tBu	OH	OCH <sub>3</sub>	CF <sub>3</sub>	Cl	Br	F
$\pi$ alif	0.50	1.68	-1.16	0.47	1.07	0.39	0.60	-0.17
$\pi$ arom	0.52	1.68	-0.67	-0.02	1.16	0.71	0.86	0.14

- Il valore di  $\pi$  è valido per strutture simili ('parenti')
- È possibile calcolare il  $\log P$  usando i valori di  $\pi$



*meta*-Chlorobenzamide

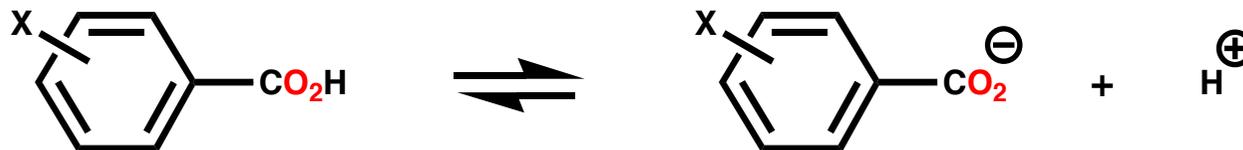
$$\begin{aligned} \text{Log } P_{(\text{theory})} &= \log P_{(\text{benzene})} + \pi_{\text{Cl}} + \pi_{\text{CONH}_2} \\ &= 2.13 + 0.71 - 1.49 = 1.35 \\ \text{Log } P_{(\text{observed})} &= 1.51 \end{aligned}$$

- Equazione QSAR: può includere sia  $P$  che  $\pi$
- $P$  è relativo all'importanza dell'idrofobicità complessiva della molecola (rilevante per assorbimento, legame, ecc)
- $\pi$  identifica specifiche regioni della molecola che potrebbero interagire con regioni idrofobiche nel sito di legame

# Effetti elettronici

## COSTANTE DI HAMMETT ( $\sigma$ )

- La costante ( $\sigma$ ) è una misura dell'influenza elettron-attrattrice/donatrice dei sostituenti
- Misurata sperimentalmente e tabulata (es  $\sigma$  per sostituenti aromatici è misurata comparando le costanti di dissociazione di acidi benzoici sostituiti con quella dell'acido benzoico)



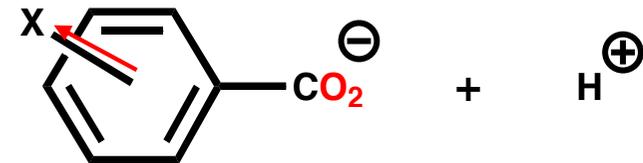
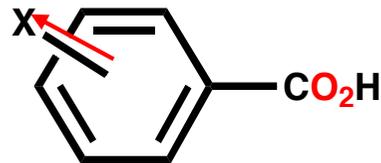
$$X = \text{H} \quad K_{\text{H}} = \text{Costante di dissociazione} = \frac{[\text{PhCO}_2^-]}{[\text{PhCO}_2\text{H}]}$$

# Effetti elettronici

## COSTANTE DI HAMMETT ( $\sigma$ )

### Gruppi elettronattrattori (e.g. NO<sub>2</sub>)

X = electron  
withdrawing  
group



Carica stabilizzata da X  
Equilibrio spostato verso destra  
 $K_X > K_H$

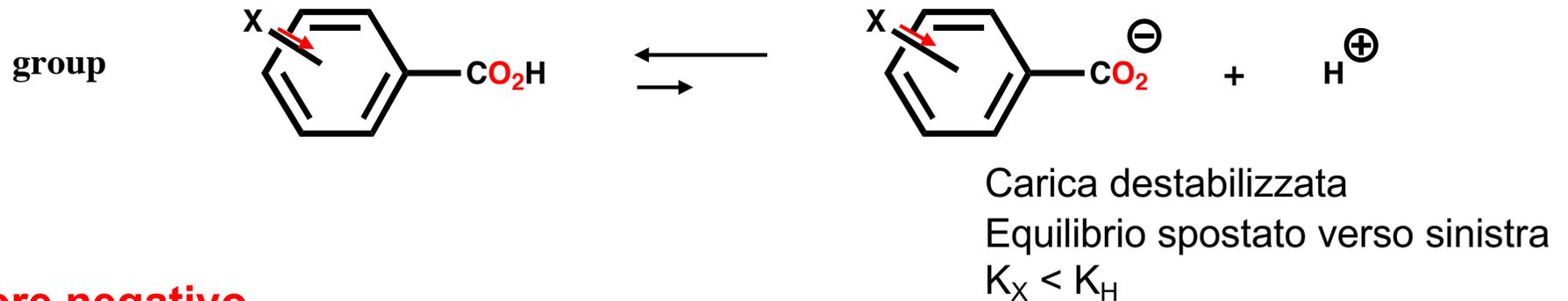
Valore positivo

$$\sigma_X = \log \frac{K_X}{K_H} = \log K_X - \log K_H$$

# Effetti elettronici

## COSTANTE DI HAMMETT ( $\sigma$ )

### Gruppi elettrondonatori (e.g. CH<sub>3</sub>)



### Valore negativo

$$\sigma_x = \log \frac{K_x}{K_H} = \log K_x - \log K_H$$

$\sigma$  dipende dagli effetti induttivo e di risonanza

$\sigma$  dipende dalla posizione del sostituente (*meta* o *para*)

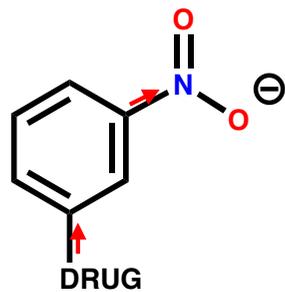
Valori dei sostituenti in *orto* non sono significativi (influenza dei fattori sterici)

# Effetti elettronici

## COSTANTE DI HAMMETT ( $\sigma$ )

e-attrattori

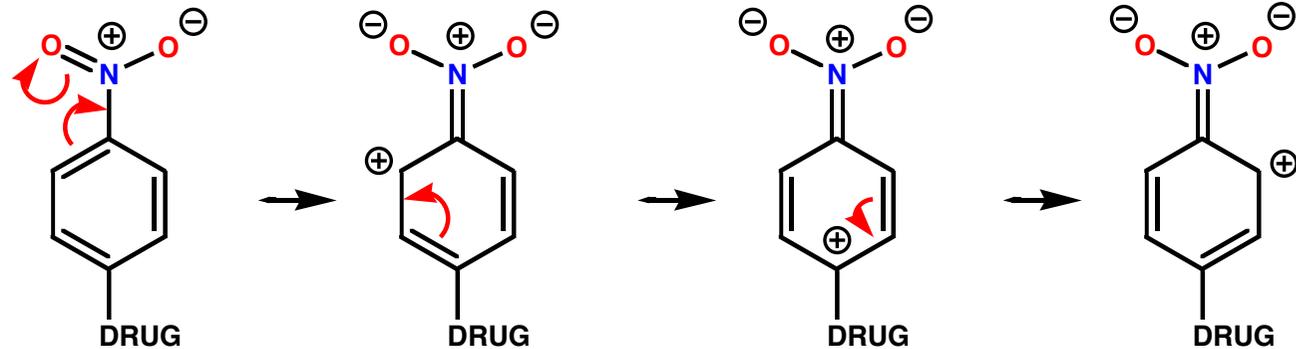
*meta*



Solo effetto induttivo

$$\sigma_m (\text{NO}_2) = 0.71$$

*para*



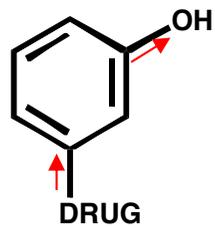
Effetto induttivo e di risonanza

$$\sigma_p (\text{NO}_2) = 0.78$$

# Effetti elettronici

## COSTANTE DI HAMMETT ( $\sigma$ )

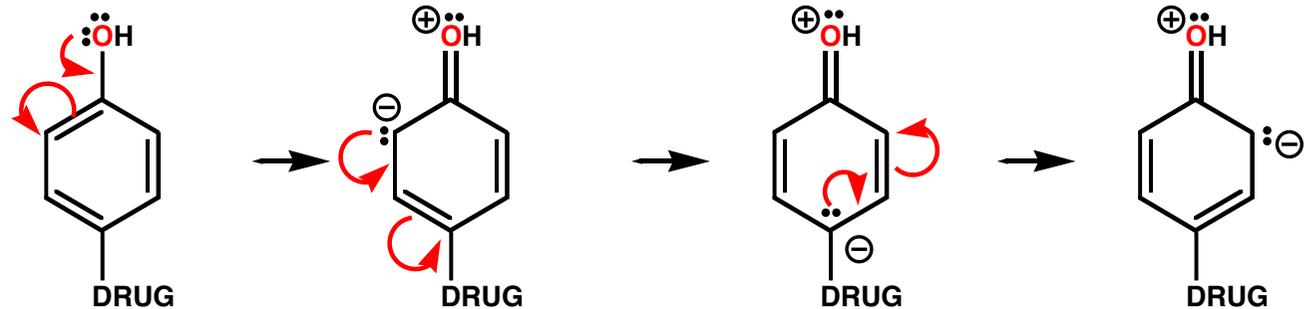
*meta*



e-attractore  
(Solo effetto induttivo)

$$\sigma_m (\text{OH}) = 0.12$$

*para*



e-donatore: effetto di risonanza più importante di quello induttivo

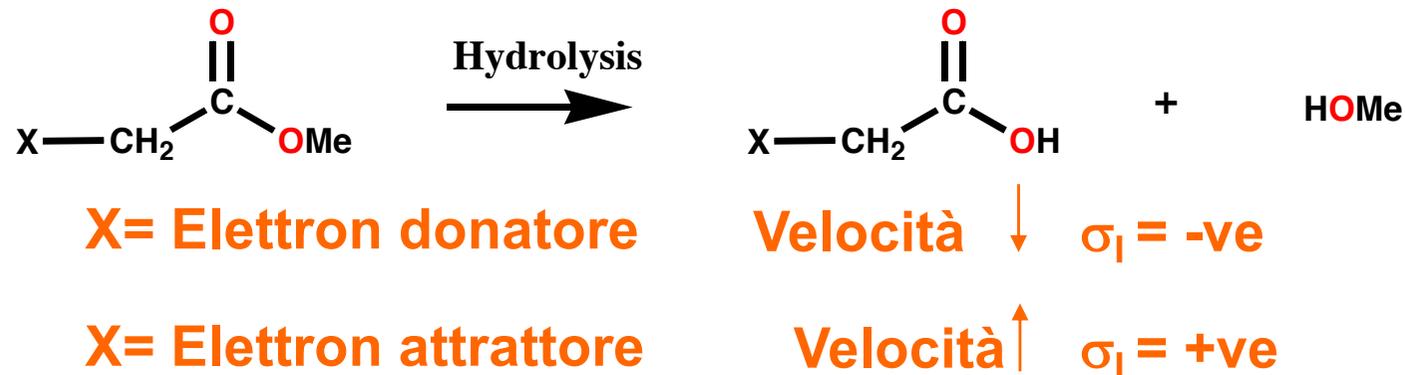
$$\sigma_p (\text{OH}) = -0.37$$

# Effetti R & F

- *R* – Quantifica gli effetti di risonanza di un sostituente
- *F* – Quantifica gli effetti induttivi di un sostituente

## Sostituenti alifatici

- Definiti da  $\sigma_I$
- Effetto puramente induttivo
- Ottenuti sperimentalmente misurando la velocità d'idrolisi degli esteri alifatici
- Velocità d'idrolisi misurata in condizioni acide e basiche



Condizioni basiche: Velocità influenzata da fattori **sterici ed elettronici**  
 Valore di  $\sigma_I$  dopo la correlazione con gli effetti sterici

Condizioni acide: Velocità influenzata solo da fattori **sterici** ( $E_s$ )

# Effetti sterici

## Fattore sterico di Taft ( $E_s$ )

- Determinato mediante confronto della **velocità d'idrolisi** di **esteri alifatici sostituiti** con quella di un **estere standard**, in **condizioni acide**

$$E_s = \log k_x - \log k_o$$

$k_x$ : velocità di idrolisi di un estere sostituito

$k_o$ : velocità di idrolisi dell'estere non sostituito

- Limitato ai sostituenti che stericamente interagiscono con lo stato di transizione della reazione
- Non può essere usato per sostituenti che interagiscono con lo stato di transizione mediante risonanza o legami idrogeno
- Possibile sottostima dell'effetto sterico dei gruppi in un processo intermolecolare (legame farmaco/recettore)

Me	0	<i>n</i> -Bu	-0.39
H	1.24	<i>i</i> -Pr	-0.47
F	0.78	<i>i</i> -Bu	-0.93
Et	-0.07	Cyclopentyl	-0.51
<i>n</i> -Pr	-0.36		

# Effetti sterici

## Refrattività molare (*MR*)

Misura del volume dei sostituenti

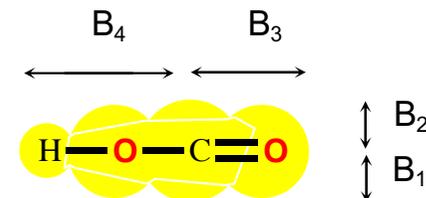
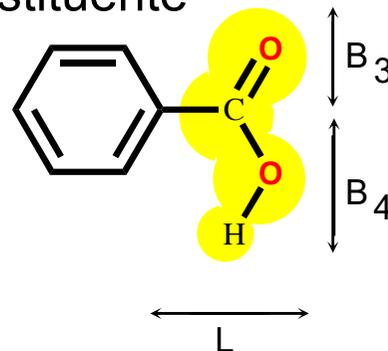
$$MR = \frac{(n^2 - 1)}{(n^2 - 2)} \times \frac{\text{mol. wt.}}{\text{density}}$$

Fattore di correzione per  
la polarizzazione  
( $n$ =indice di rifrazione)

Definisce il volume

- calcolato mediante software (STERIMOL)
- dà le dimensioni di un sostituente
- può essere utilizzato per qualsiasi sostituente

## Parametro sterico di Verloop

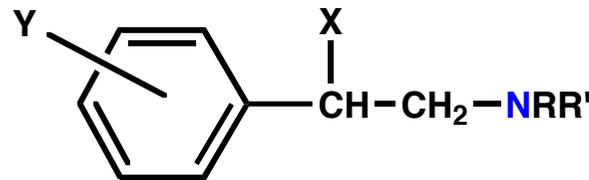


# Equazione di Hansch

- Equazione QSAR che correla varie **proprietà chimico-fisiche** con l'attività biologica di una serie di composti
- Di solito include logP, fattori elettronici e sterici
- Inizia con semplici equazioni che vengono maggiormente elaborate con i dati ottenuti da un numero sempre maggiore di composti sintetizzati
- Equazione tipica per una vasta gamma di logP: parabolica

$$\text{Log } (1/C) = -k_1 (\log P)^2 + k_2 \log P + k_3 \pi + k_4 \sigma + k_5 E_S + k_6$$

Attività antiadrenergica di  $\beta$ -  
alo- $\beta$ -arilammine

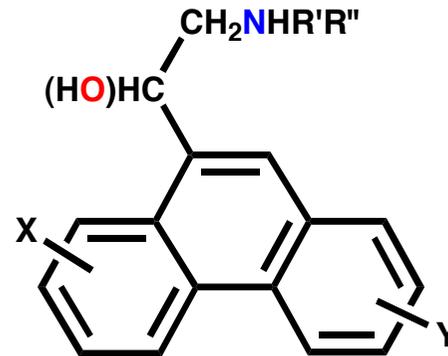


$$\text{Log } (1/C) = 1.22 \pi - 1.59 \sigma + 7.89$$

- Aumento di attività se  $\pi$  positivo (i.e. sostituenti idrofobici)
- Aumento di attività se  $\sigma$  negativo (i.e. sostituenti e-donatori)

# Equazione di Hansch

Attività antimalarica di amminocarbinoli fenantrenici



$$\text{Log (1/C)} = -0.015 (\text{log P})^2 + 0.14 \text{ log P} + 0.27 \Sigma\pi_X + 0.40 \Sigma\pi_Y + 0.65 \Sigma\sigma_X + 0.88 \Sigma\sigma_Y + 2.34$$

- Attività lievemente maggiore con l'aumento del logP (idrofobicità) aumenta (la costante è soltanto 0.14!)
- Equazione parabolica: valore ottimale di logP per l'attività ( $P^0$ )
- Attività aumentata con sostituenti idrofobici (Anello Y)
- Attività aumentata con sostituenti e-attrattori (esp. Anello di Y)

# Equazione di Hansch

## Sostituenti adatti

I sostituenti devono essere scelti in modo tale da soddisfare i seguenti criteri:

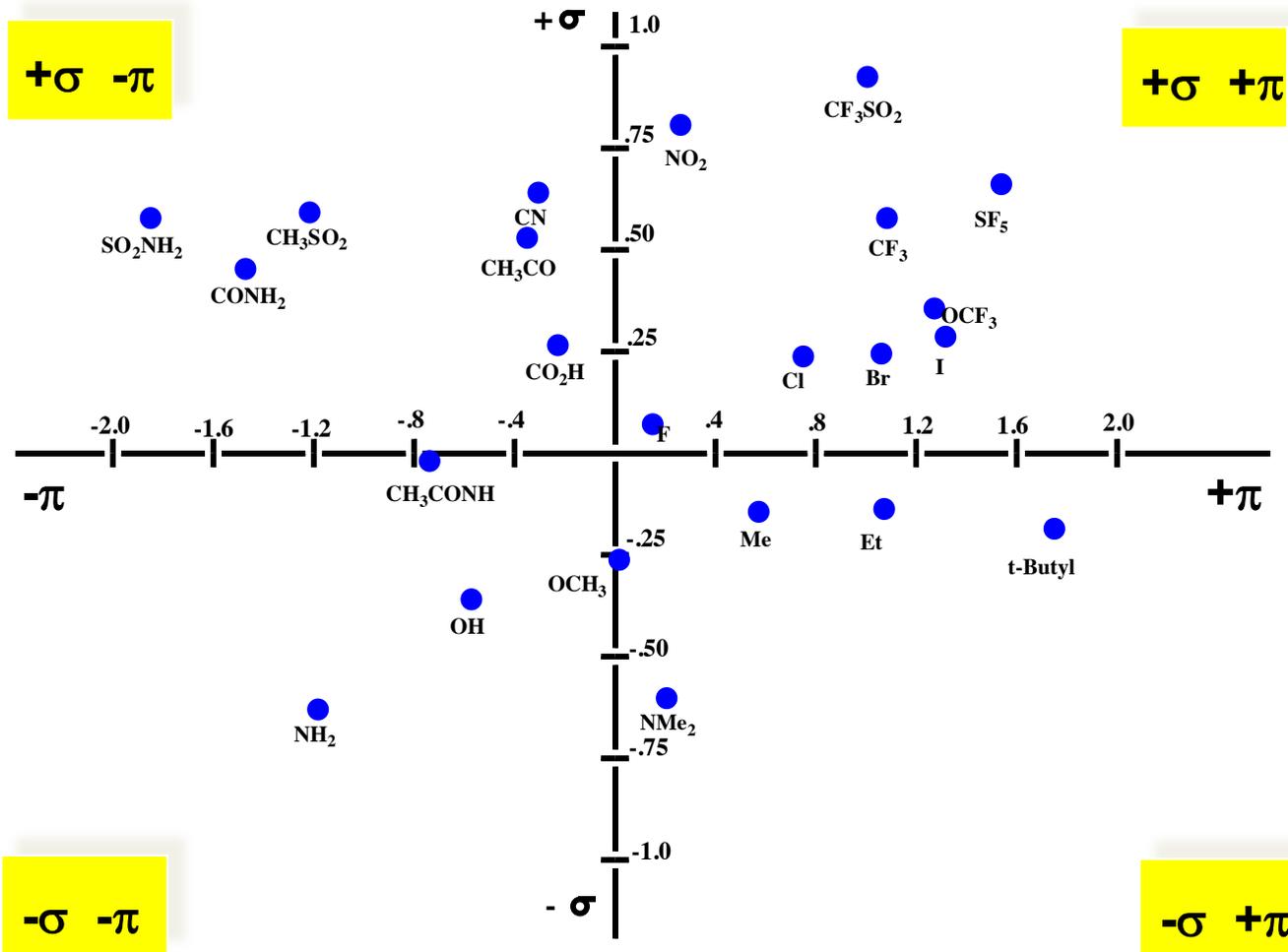
- Intervallo di valori per ogni proprietà chimico-fisica studiata
- Valori non correlati per le diverse proprietà (ortogonali)
- Almeno 5 strutture per ciascun parametro studiato

Sostituente	H	Me	Et	n-Pr	n-Bu	} Valori correlati Differenze dovute a $\pi$ o MR?
$\pi$	0.00	0.56	1.02	1.50	2.13	
MR	0.10	0.56	1.03	1.55	1.96	

Sostituente	H	Me	OMe	NHCONH <sub>2</sub>	I	CN	} Valori non correlati Valido per analizzare gli effetti di $\pi$ e MR
$\pi$	0.00	0.56	-0.02	-1.30	1.12	-0.57	
MR	0.10	0.56	0.79	1.37	1.39	0.63	

# Plot di Craig

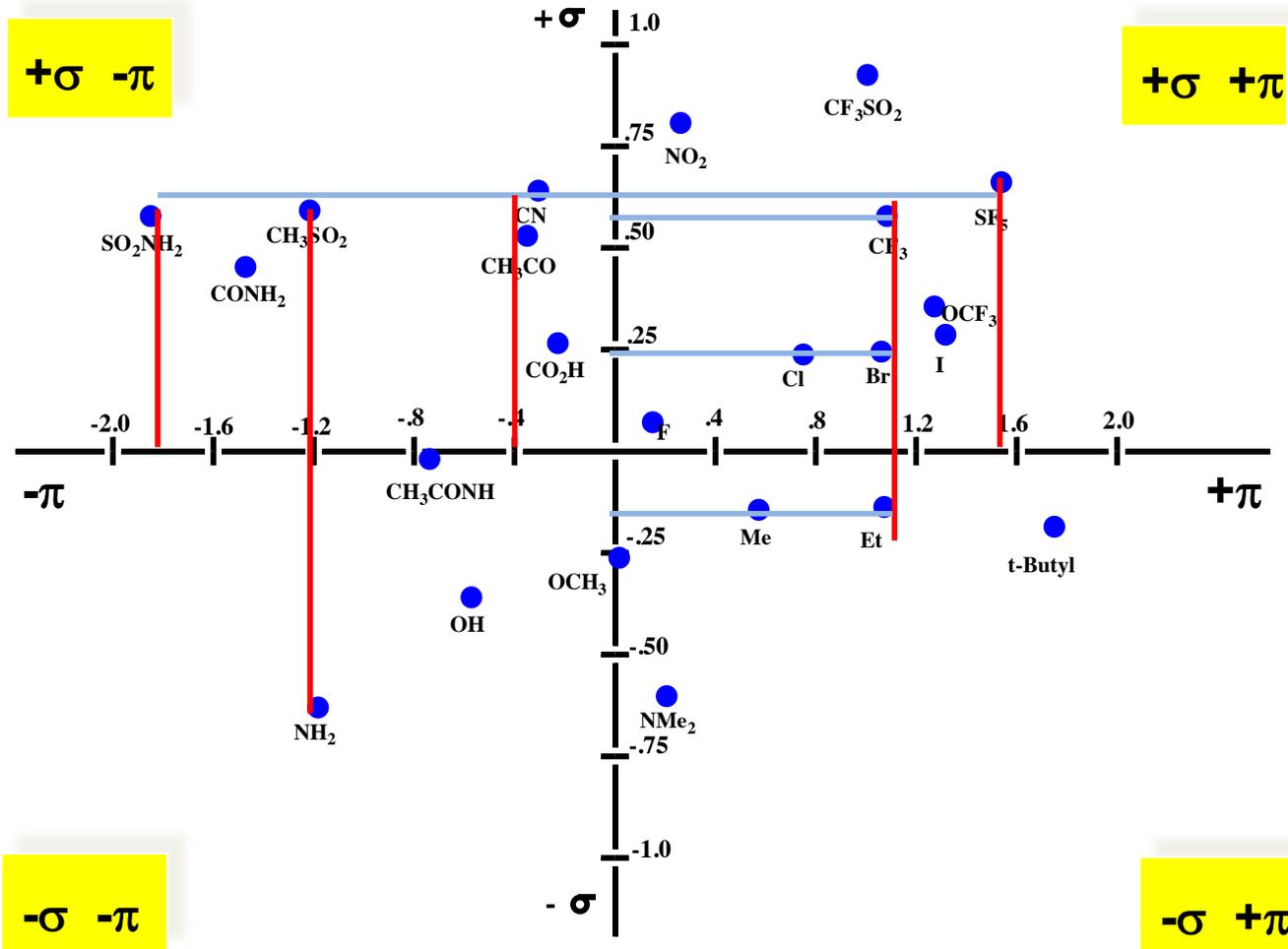
Mostra i valori di 2 diverse proprietà chimicofisiche per molti sostituenti



- Permette una facile identificazione dei sostituenti adatti per un'analisi QSAR che comprenda entrambe le proprietà
- Scelta di un sostituito da ogni quadrante per garantire ortogonalità
- Scelta di sostituenti con un intervallo di valori per ogni proprietà

# Plot di Craig

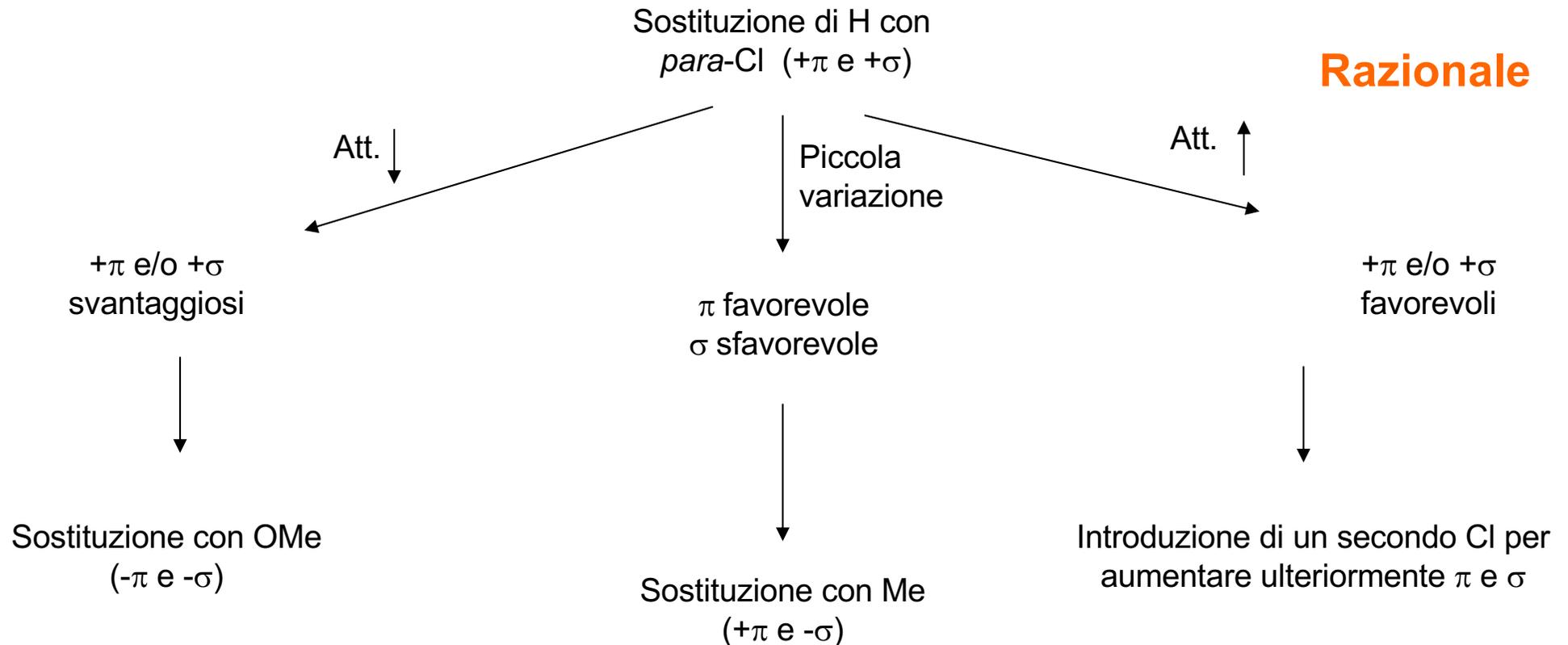
Mostra i valori di 2 diverse proprietà chimicofisiche per molti sostituenti



- Permette una facile identificazione dei sostituenti adatti per un'analisi QSAR che comprenda entrambe le proprietà
- Scelta di un sostituito da ogni quadrante per garantire ortogonalità
- Scelta di sostituenti con un intervallo di valori per ogni proprietà

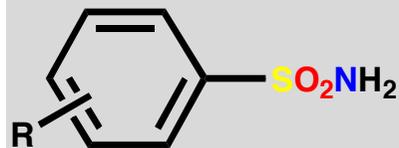
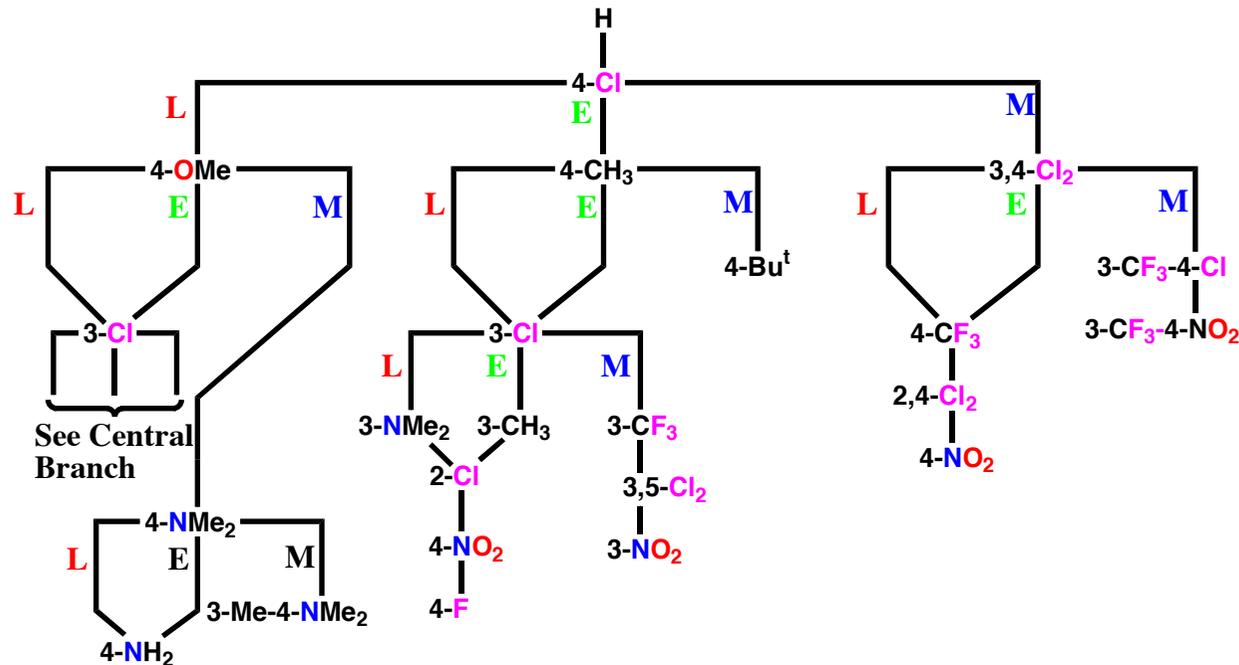


# Schema di Topliss



Ulteriori cambiamenti suggeriti sulla base di  $\pi$ ,  $\sigma$  e dell'ingombro sterico

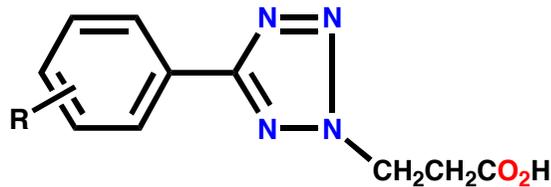
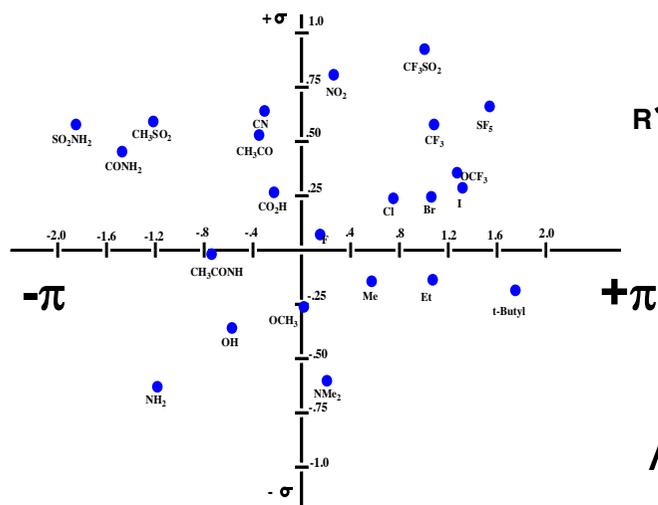
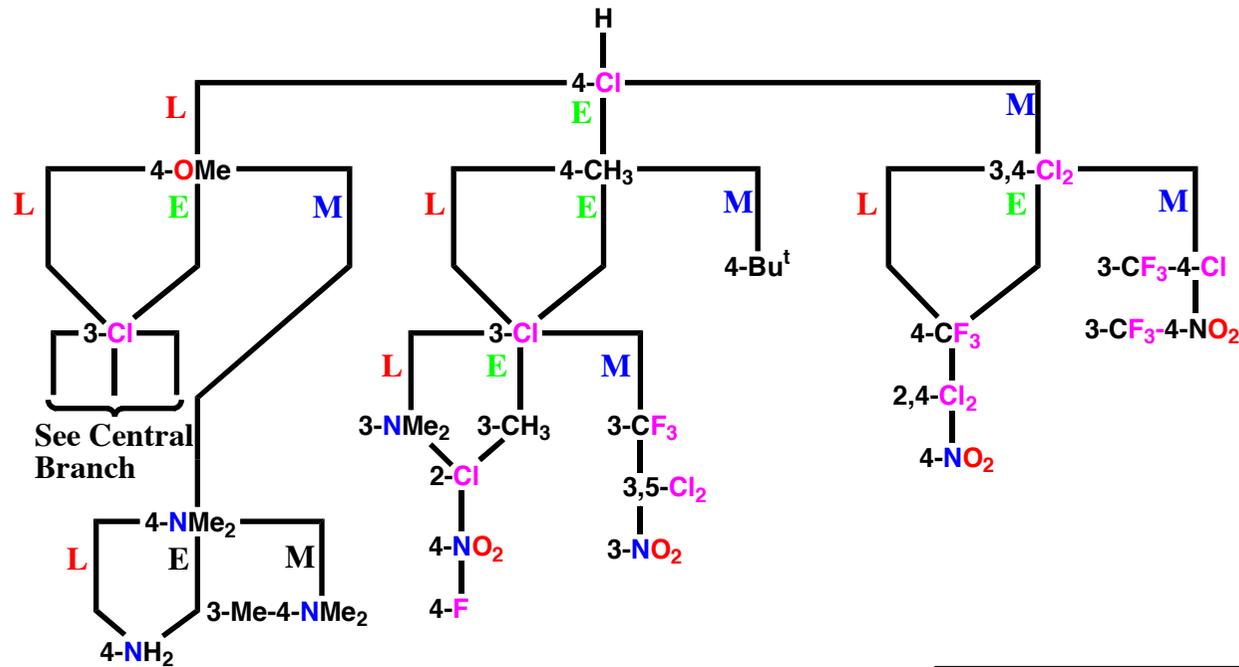
# Schema di Topliss



Order of Synthesis	R	Biological Activity	High Potency
1	H	-	
2	4-Cl	M	
3	3,4-Cl <sub>2</sub>	L	
4	4-Br	E	
5	4-NO <sub>2</sub>	M	*

M= More Activity  
 L= Less Activity  
 E = Equal Activity

# Schema di Topliss

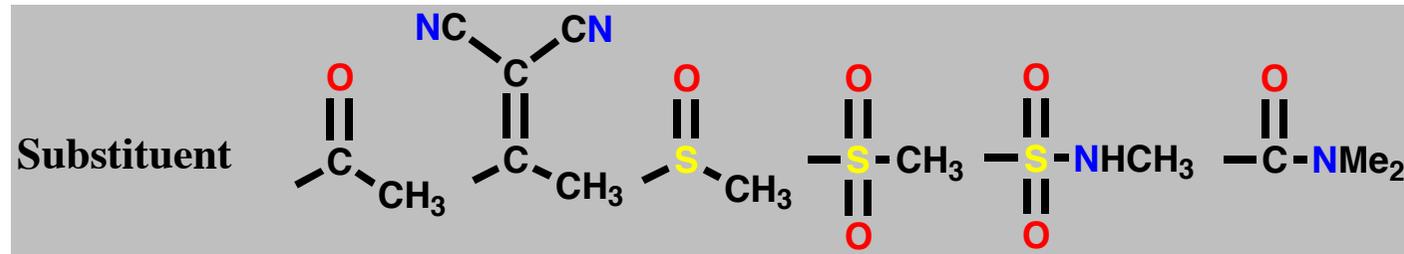


Attività antiinfiammatoria

Order of Synthesis	R	Biological Activity	High Potency
1	H	-	
2	4-Cl	L	
3	4-MeO	L	
4	3-Cl	M	*
5	3-CF <sub>3</sub>	L	
6	3-Br	M	*
7	3-I	L	
8	3,5-Cl <sub>2</sub>	M	*

M= More Activity  
L= Less Activity  
E = Equal Activity

# Sostituenti isosterici/bioisosterici

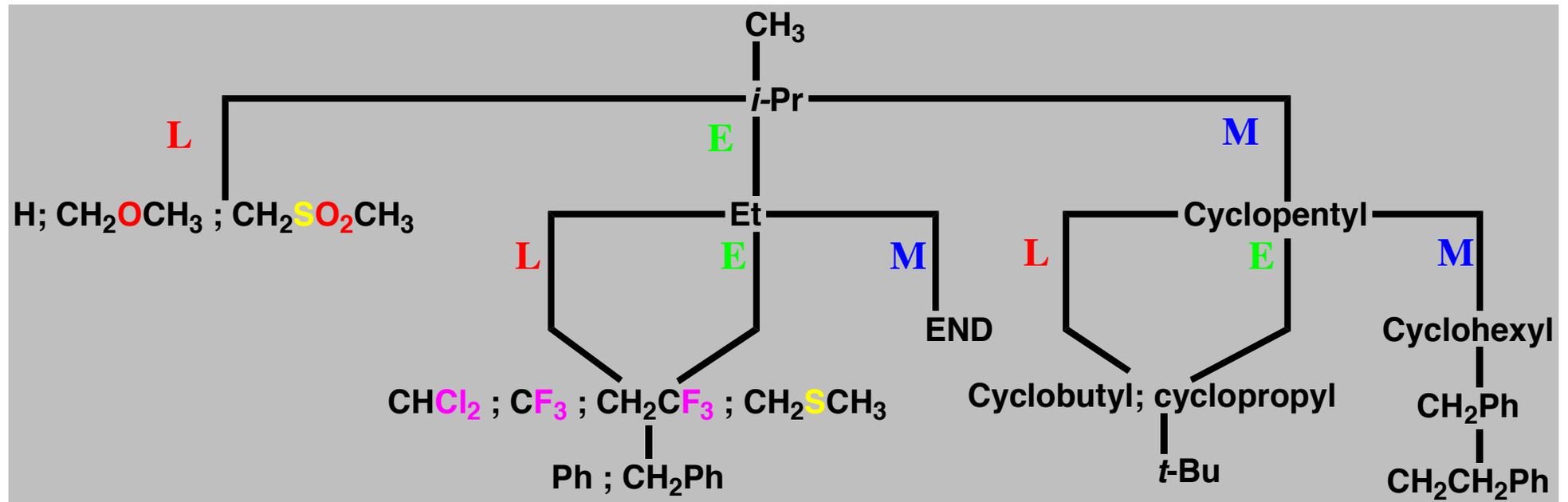


$\pi$	-0.55	0.40	-1.58	-1.63	-1.82	-1.51
$\sigma_p$	0.50	0.84	0.49	0.72	0.57	0.36
$\sigma_m$	0.38	0.66	0.52	0.60	0.46	0.35
<i>MR</i>	11.2	21.5	13.7	13.5	16.9	19.2

- Scelta di sostituenti con simili proprietà fisico-chimiche (i.e. CN, NO<sub>2</sub> e COMe potrebbero essere bio-isosteri)
- Scelta di bio-isosteri sulla base delle proprietà chimico-fisiche più importante (i.e. COMe e SOMe simili per  $\sigma_p$ ; COMe e SO<sub>2</sub>Me presentano simile valore di  $\pi$ )

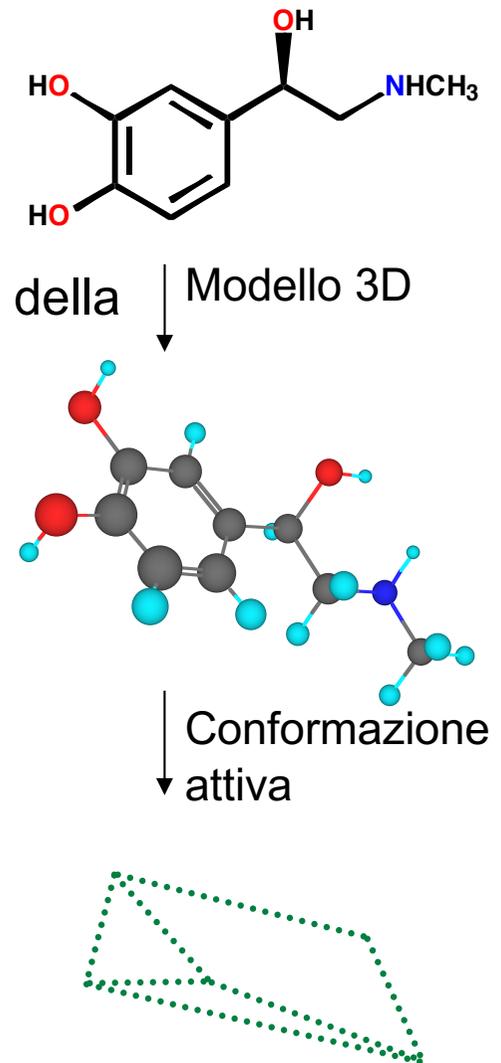
# Schema di Topliss

## Sostituenti alifatici



# 3D- QSAR

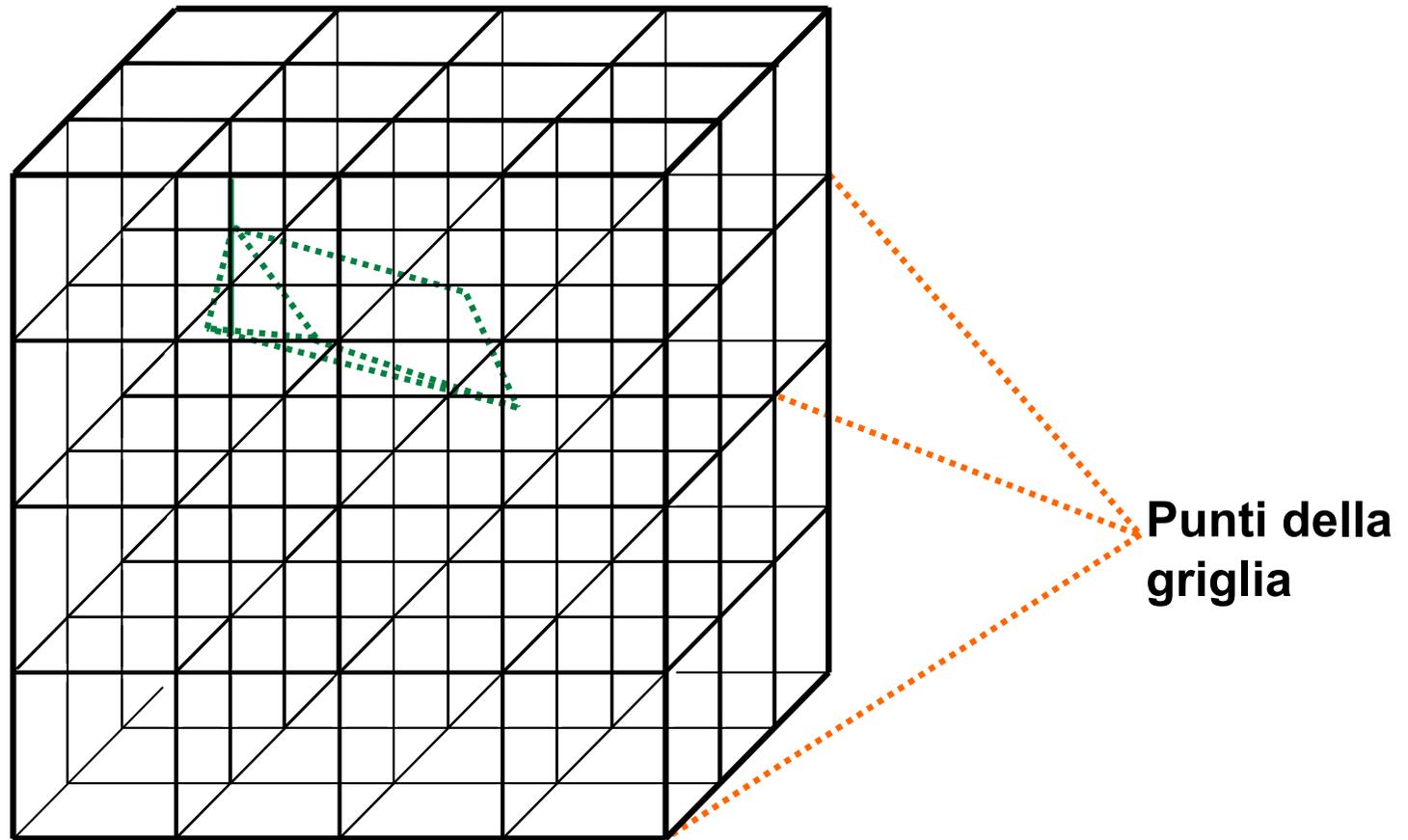
- Proprietà fisiche misurate per la molecola nel suo complesso
- Proprietà calcolate utilizzando software computazionali
- NO costanti o misure sperimentali
- Proprietà identificate come 'campi'
- Campo sterico – definisce dimensione e forma della molecola
- Campo elettrostatico - definisce regioni elettrone-ricche / povere della molecola
- Proprietà idrofobe relativamente poco importanti
- Nessun affidamento su valori sperimentali
- Applicabile a molecole con sostituenti inusuali
- Non limitato alle molecole della stessa classe strutturale
- Capacità predittiva
- Analisi comparativa dei campi molecolari  
Comparative molecular field analysis (CoMFA) – Tripos
- “Costruzione” della molecola in silico
- Identificazione della conformazione attiva di ogni molecola
- Identificazione del farmacoforo



Definizione del farmacoforo

# 3D- QSAR

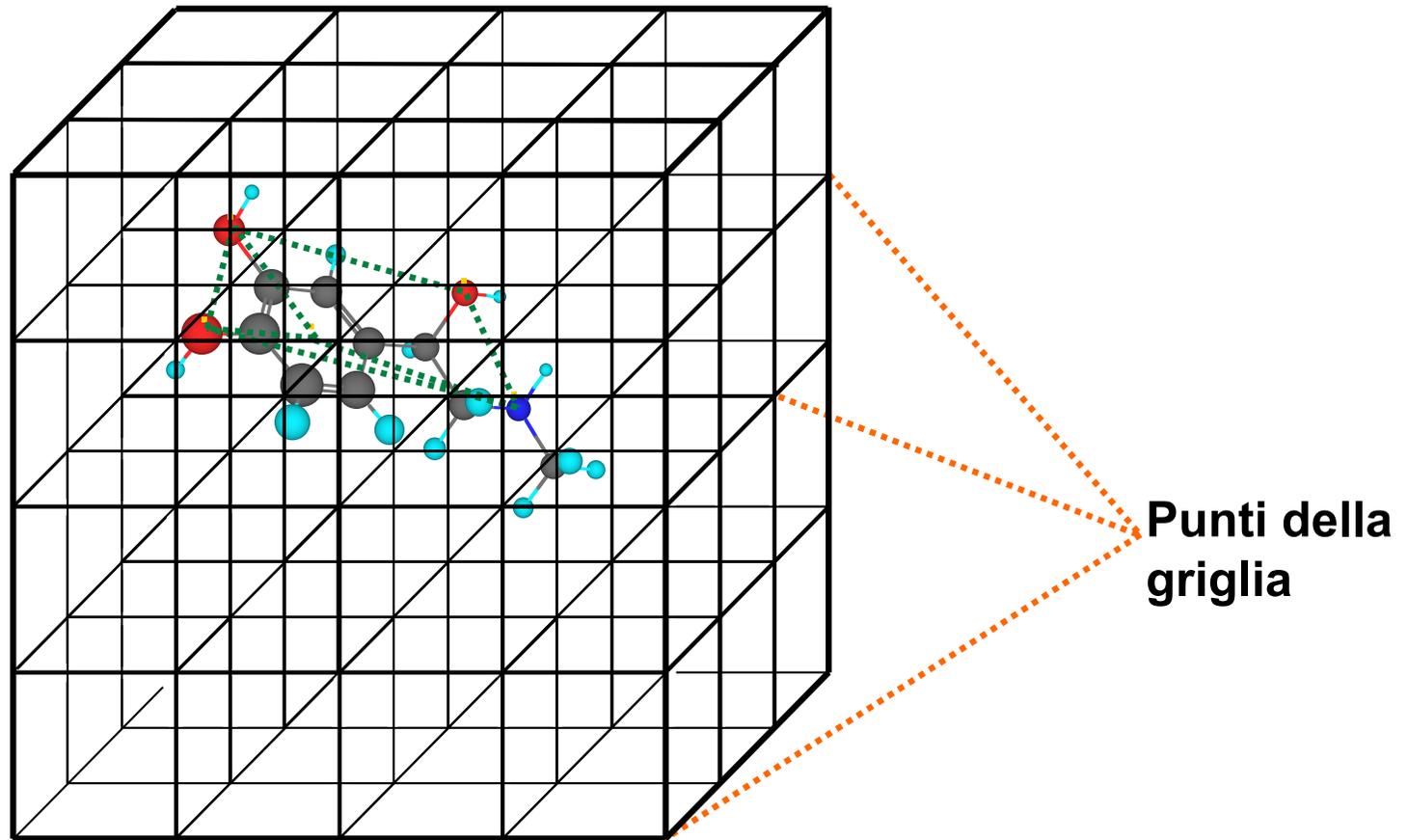
Posizionamento della molecola in modo che combaci con il farmacoforo



Ciascun punto della griglia definisce un punto dello spazio

# 3D- QSAR

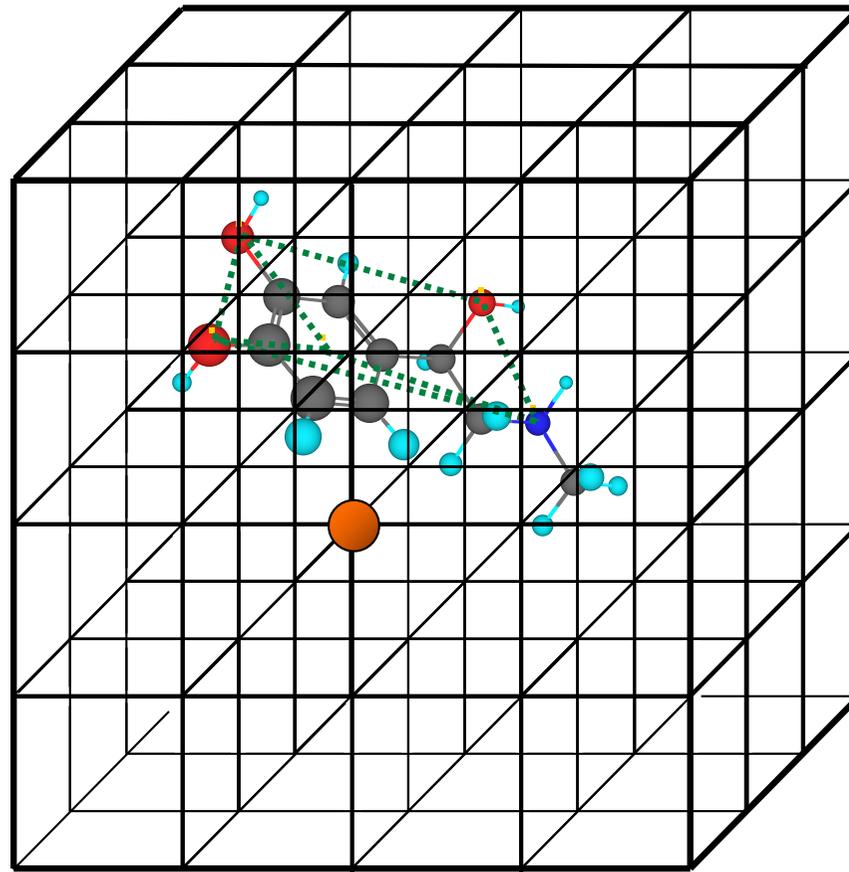
Posizionamento della molecola in modo che combaci con il farmacoforo



Ciascun punto della griglia definisce un punto dello spazio

# 3D- QSAR

Un atomo sonda è viene spostato su ciascun punto della griglia

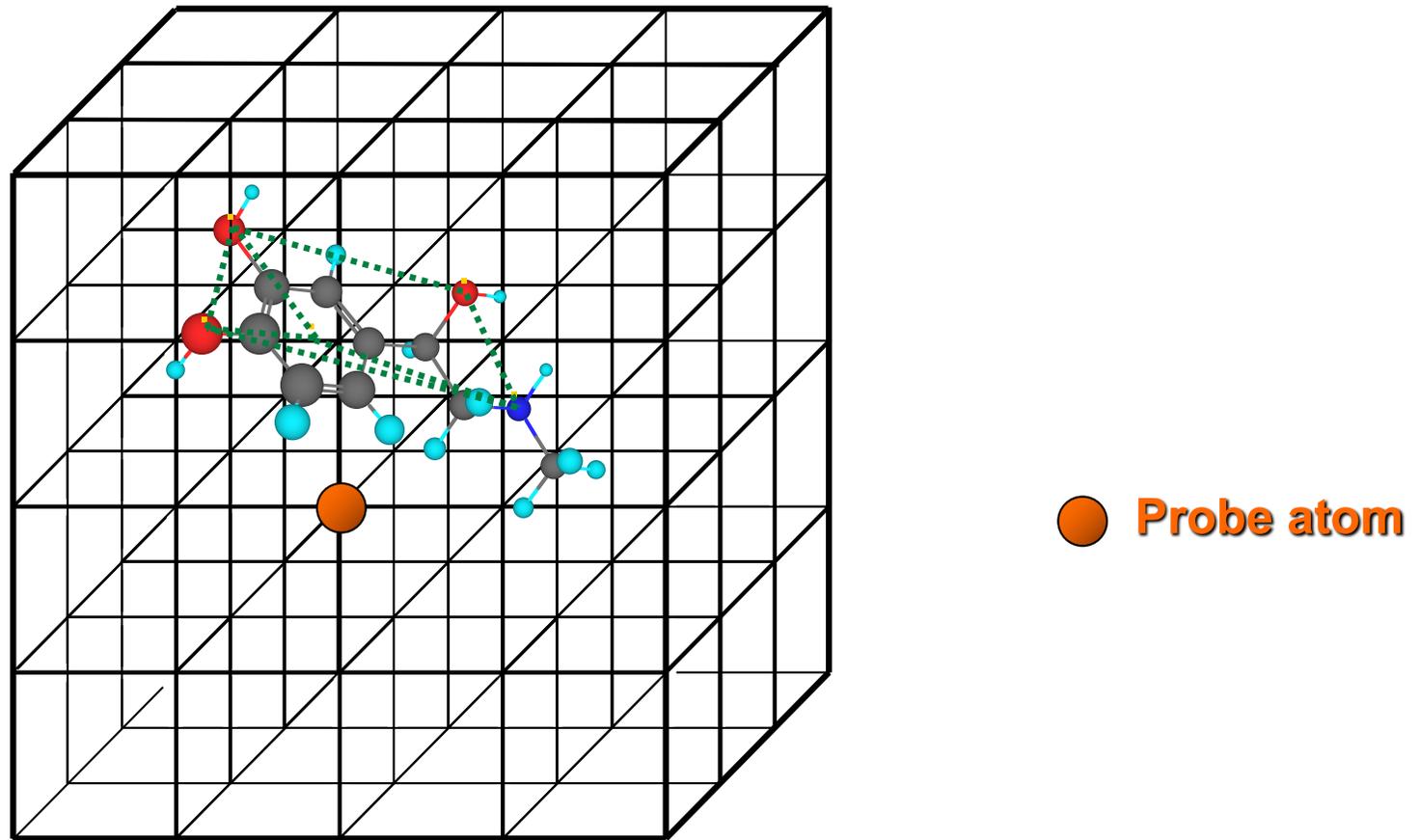


● Probe atom

Atomo sonda = un protone o un carbocatione ibridizzato  $sp^3$

# 3D- QSAR

Un atomo sonda è viene spostato su ciascun punto della griglia

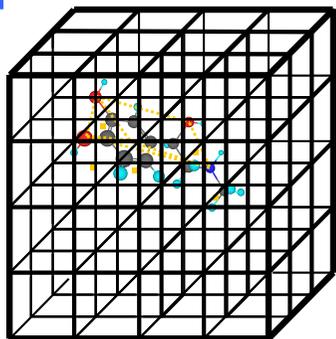


Misura dell'interazione sterica o elettrostatica dell'atomo sonda con la molecola in ciascun punto della griglia

## 3D- QSAR

- Quanto più l'atomo sonda è vicino alla molecola, tanto maggiore è l'energia sterica
- Definizione della forma della molecola, individuando punti della griglia con uguale energia sterica (linee di contorno)
- Interazioni elettrostatiche favorevoli evidenziate con la sonda di carica positiva: regioni molecolari che presentano negativa
- Interazioni elettrostatiche sfavorevoli evidenziate con la sonda di carica positiva: regioni molecolari con carica positiva
- Definizione di campi elettrostatici, individuando punti di griglia di uguale energia (linea di contorno)
- Procedura ripetuta per ogni molecola da analizzare
- Confronto dei campi ottenuti per ciascuna molecola con l'attività biologica
- Identificazione dei campi sterici ed elettrostatici che sono favorevoli o sfavorevoli per l'attività

# 3D- QSAR



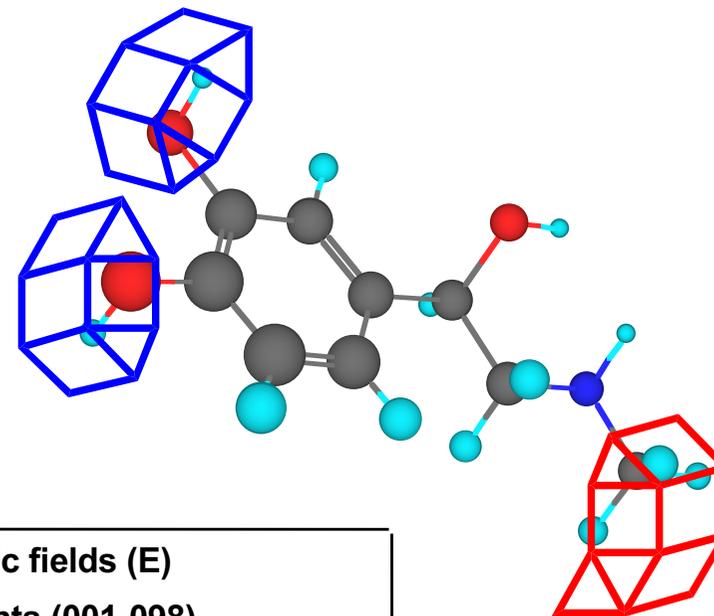
↓  
Tabulate fields for each  
compound at each grid point

Compound	Biological activity	Steric fields (S) at grid points (001-998)					Electrostatic fields (E) at grid points (001-098)				
		S001	S002	S003	S004	S005 etc	E001	E002	E003	E004	E005 etc
1	5.1										
2	6.8										
3	5.3										
4	6.4										
5	6.1										

↓  
Partial least squares  
analysis (PLS)

QSAR equation      Activity = aS001 + bS002 + .....mS998 + nE001 + .....+yE998 + z

CROSS VALIDATION → Cross validated correlation coefficient  $q^2$  (>0.3)

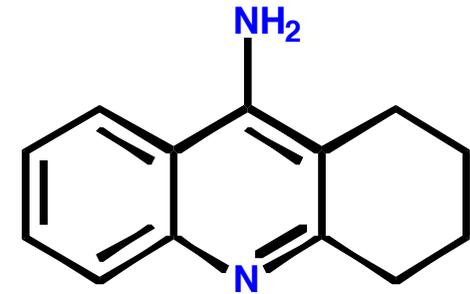


Definizione dei  
campi usando  
mappe colorate  
intorno a una  
molecola  
rappresentativa

# 3D- QSAR

## Tacrina

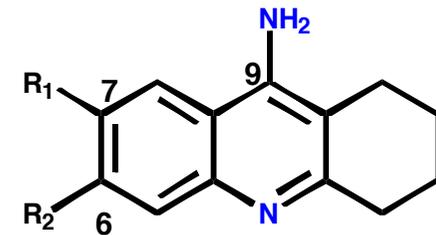
Inibitore dell'acetilcolinesterasi usato nel trattamento del morbo d'Alzheimer



## Studi QSAR convenzionali

12 analoghi la cui attività è stata correlata alle proprietà idrofobiche, steriche ed elettroniche dei sostituenti in 6 e 7

Substituenti: CH<sub>3</sub>, Cl, NO<sub>2</sub>, OCH<sub>3</sub>, NH<sub>2</sub>, F  
(Valori 'dispersi' senza correlazione)



$$\text{Log}\left(\frac{1}{C}\right) = p \text{ IC}_{50} = - 3.09 \text{ MR}(R_1) + 1.43 \text{ F}(R_1, R_2) + 7.00$$

*F* – Quantifica l'effetto induttivo dei sostituenti

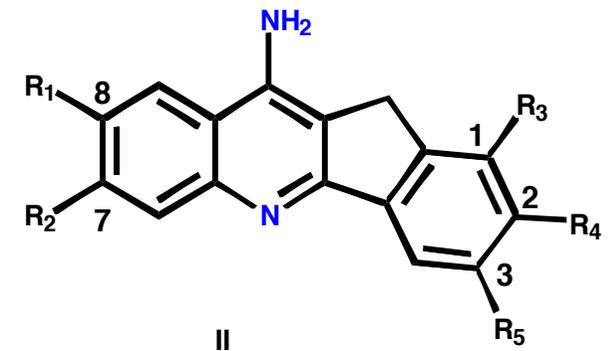
Gruppi ingombranti in posizione 7: diminuzione dell'attività  
Gruppi elettron-attrattori in posizione 6 & 7  
Non c'è effetto idrofobico

# 3D- QSAR

## Studi CoMFA

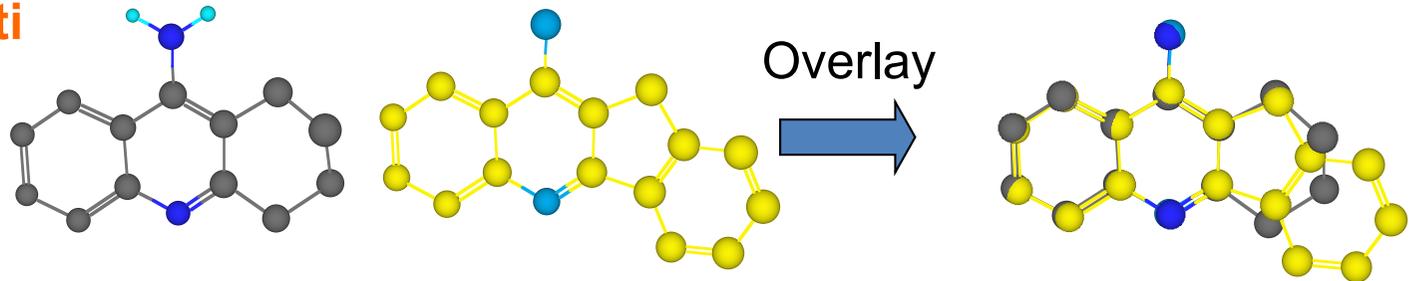
### Analisi di inibitori anticolinesterasici tetraciclici (II)

- Non è possibile includere queste strutture in un'analisi QSAR convenzionale poiché appartengono a una classe strutturale diversa
- Molecole appartenenti a classi strutturali diverse devono essere allineate correttamente secondo un modello farmacoforico comune

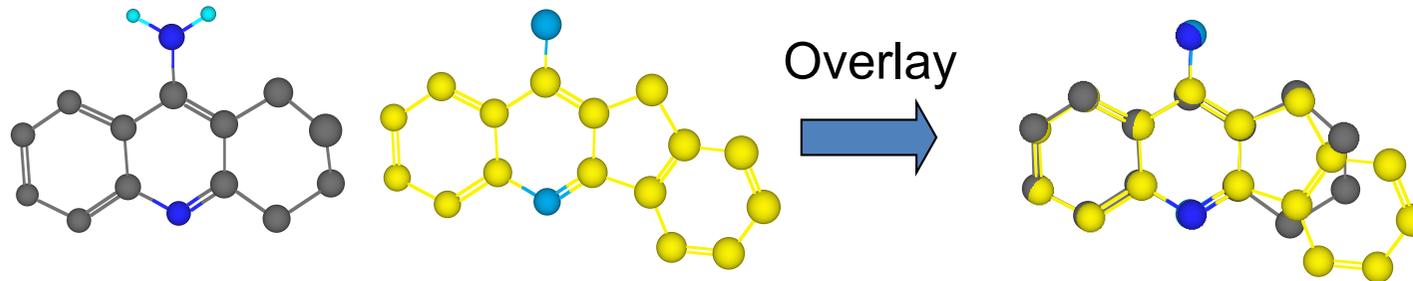


### Possibili allineamenti

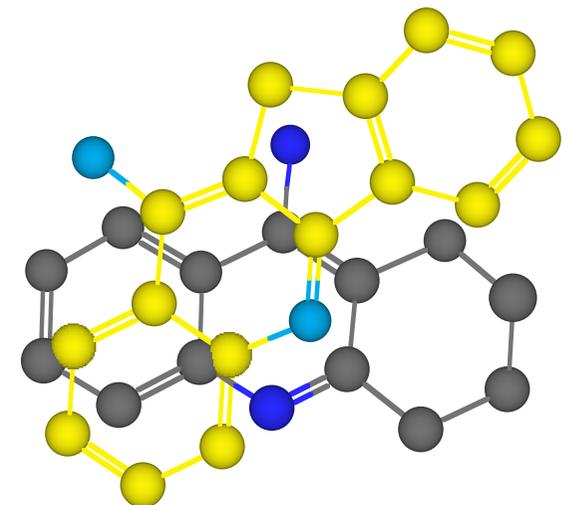
Buona sovrapposizione ma si ASSUME che la modalità di binding sia la stessa



# 3D- QSAR



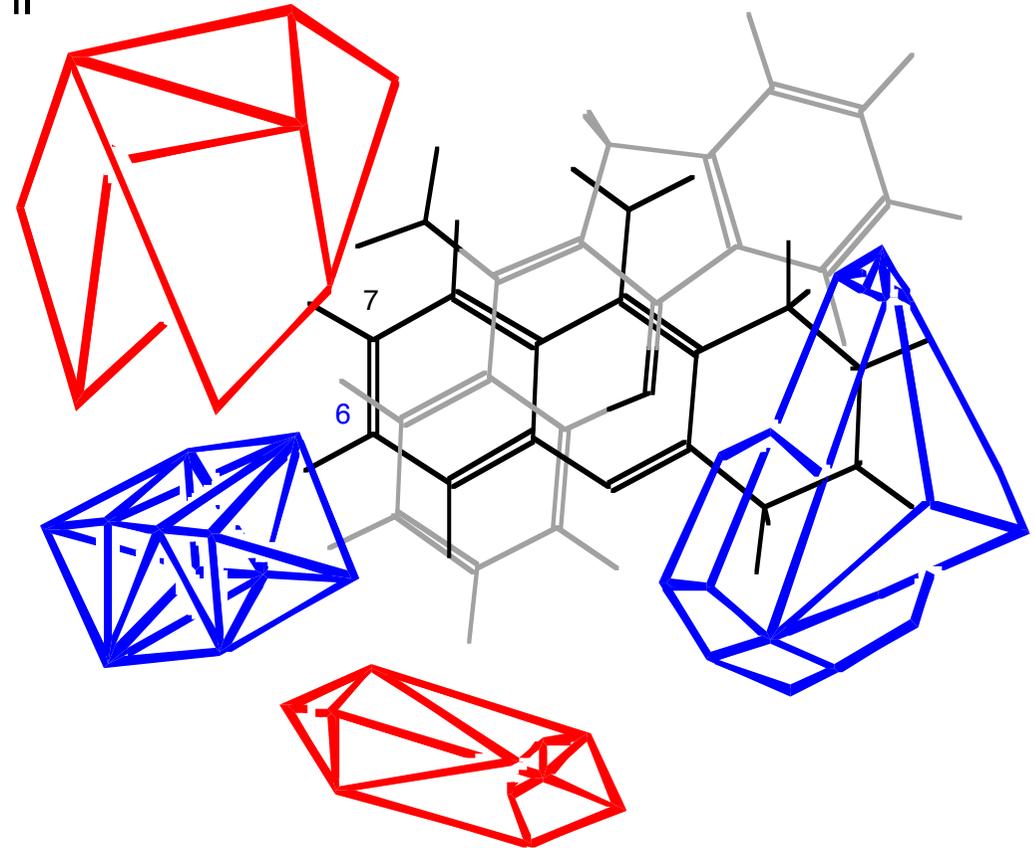
- Complesso tacrina/enzima cristallizzato e analizzato
- Risultati che dimostrano il modo in cui si lega la tacrina
- Modellistica molecolare: utilizzata per modificare la tacrina nella struttura (II) quando legata al sito di legame (in silico)
- Studi computazionali sul struttura (II)/complesso (minimizzazione)
- Legame di (II) diverso da quello della tacrina



# 3D- QSAR

## Possibile allineamento

- Analoghi di ogni tipo di struttura allineati secondo la struttura "parente"
- L'analisi mostra che il fattore sterico è il solo responsabile per l'attività



Area blu – aumento dell'ingombro sterico con aumento di attività

Area rosse – aumento dell'ingombro sterico diminuisce l'attività