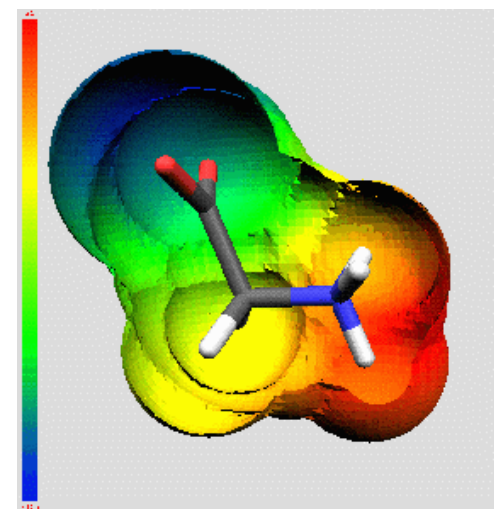


# Berechnung von Moleküleigenschaften (1)

Prinzipiell lassen sich alle Moleküleigenschaften *direkt* berechnen, die sich aus der Wellenfunktion ableiten lassen, z.B.

- Ionisationspotential (IP)
- Elektrische Multipole (Ladung, Dipolmoment, Quadrupolmoment...)
- Molekulares Elektrostatisches Potential (MEP)
- Polarisierbarkeit und Hyperpolarisierbarkeit
- Schwingungsspektrum (IR, Raman)
- Kernspinresonanz (NMR)
- Hyperfeinkopplungskonstanten (EPR)



Damit können experimentelle Meßgrößen in der Analytik vorhergesagt werden.

## Berechnung von Moleküleigenschaften (2)

Weitere komplexe Eigenschaften lassen sich *indirekt* durch Kalibrierung an Datensets mittels Regressionsgleichungen, SVM, neuronalen Netze oder anderen wissensbasierter Methoden ableiten.

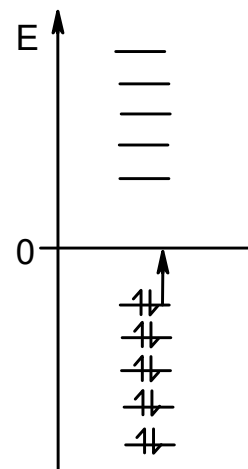
→ (Quantum) QSAR, neuronale Netze, ...

Beispiele aus der Literatur:

- logP (Wasser/n-Octanol Verteilungskoeffizient)
- $^{13}\text{C}$  NMR chemische Verschiebungen
- Mutagenizität von Verbindungen
- Löslichkeit in Wasser (logS)
- Siedepunkte
- Permeation der Blut-Hirn-Schranke (blood-brain barrier, logBB)

# Ionisationspotential IP (1)

Energie die aufzuwenden ist, um ein Elektron aus einem Orbital zu entfernen. In der Regel ist dies ein Valenzorbital (Photoelektronen Spektroskopie PES). Die Ionisierungsenergien für Elektronen aus abgeschlossenen Schalen sind wesentlich höher (X-Ray Photon Spectrum, XPS).

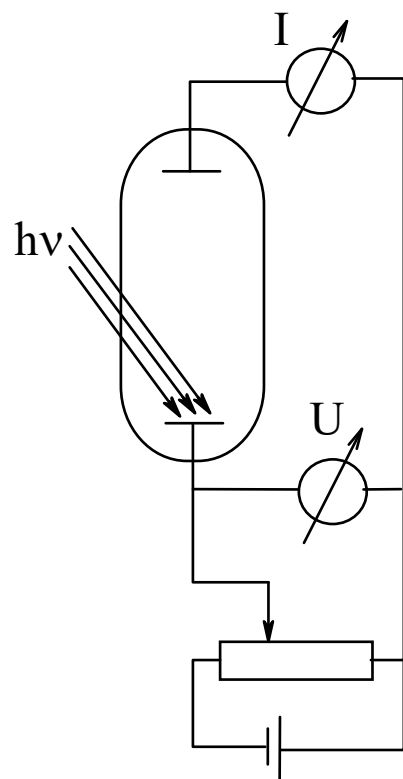


Koopmans' Theorem:  $E_{\text{ionisation}} = \epsilon_i$

Annahme: *frozen orbitals* → Orbitale behalten ihre Energien im angeregten Zustand = vertikaler Übergang, keine Änderung der Molekülgeometrie

## Ionisationspotential (2)

Experimentell: Photoelektronenspektroskopie (PES) für Valenzelektronen und Röntgen-Photoelektronenspektroskopie (XPS) für innere Elektronen



$$h\nu = E_{\text{ionisation}} + \frac{1}{2}mv_e^2$$

# Elektrische Multipole (1)

Die elektrischen Multipole geben die Ladungsverteilung im Molekül wieder:

Cl<sup>-</sup>                      Ion                      Monopol

H<sup>δ+</sup>- Cl<sup>δ-</sup>                      >=2 Ladungen      Dipol

allgemeines Dipolmoment       $\mu = \sum q_i r_i$       vektorielle Größe

Quantenmechanisch gesehen enthält das Dipolmoment Anteile der Atomkerne und der Elektronen. Während sich der Kernanteil über die Koordinaten der Atome und deren Ladung Z berechnen läßt, ergibt sich der elektronische Anteil aus der Dichtematrix **D** und den Ein-Elektronenintegralen mit dem Dipolmomentoperator R:

$$\mu_{elec} = \sum_{\mu=1}^{N_{AO}} \sum_{\nu=1}^{N_{AO}} D_{\mu\nu} \int \phi_{\mu} - R\phi_{\nu} d\tau$$

## Elektrische Multipole (2)

Die höheren elektrischen Momente sind:

CO <sub>2</sub>	4 Ladungen	Quadrupol	3*3 Matrix (Tensor)
CH <sub>4</sub>	8 Ladungen	Octupol	
	16 Ladungen	Hexadekapol ...	

Dipolmoment von Formaldehyd H<sub>2</sub>C=O (in Debye)

STO-3G	1.5258	
3-21G	2.2903	Die berechneten Dipolmomente von sehr kleinen Molekülen weisen oftmals große Abweichung zu den experimentellen Werten auf.
4-31G	3.0041	
6-31G*	2.7600	
6-311G(d,p)	2.7807	
exptl.	2.34	

# Molekulares Elektrostatisches Potential (1)

Durch die Kerne  $Z$  und Elektronen  $i$  eines Moleküls entsteht eine Ladungsverteilung im Raum. An jedem beliebigen Punkt  $r$  kann man das dadurch entstehende Potential  $V(r)$  bestimmen:

$$V_{ESP}(r) = \sum_A^{\text{Kerne}} \frac{Z_A}{|r - R_A|} - \int \frac{|\Psi(r_i)|^2}{|r - r_i|} dr_i$$

Während der Kernanteil lediglich die Ladungen  $Z$  der Kerne enthält, ist für den elektronischen Teil eine Wellenfunktion notwendig.

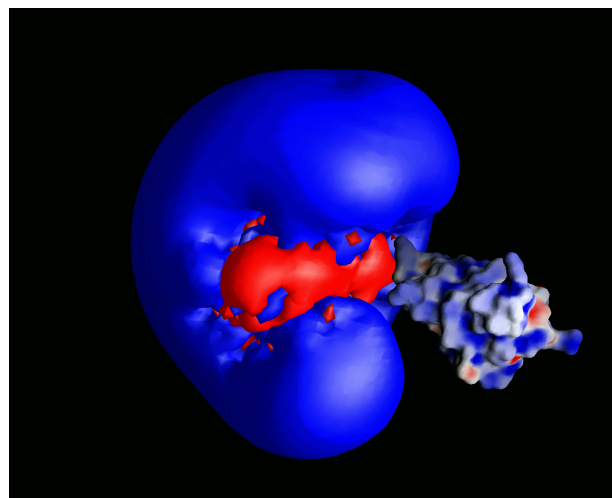
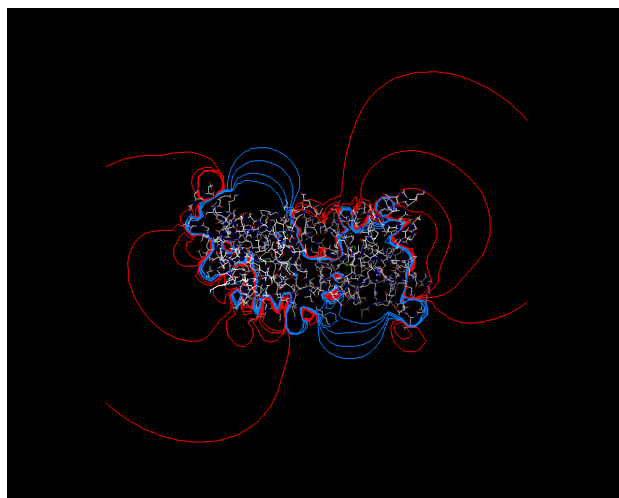
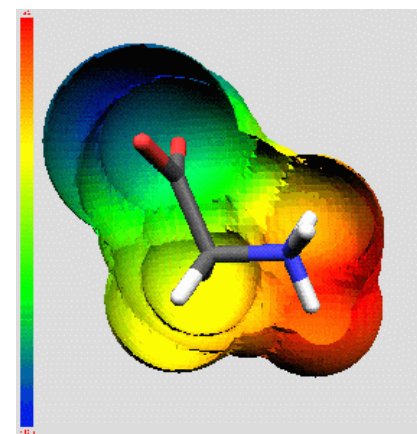
Zur Erinnerung: In Kraftfeldern benutzt man Atomladungen (auf den Atomen) die zusammenaddiert die elektrischen Multipole wiedergeben.

## Molekulares Elektrostatisches Potential (2)

Zur Bestimmung des MEP an einem Punkt  $r$  ersetzt man in der Praxis die Integration durch eine Summation über hinreichend kleine Volumenelemente.

Zur Visualisierung gibt man das MEP beispielsweise auf der van der Waals Oberfläche an.

Eine weitere Möglichkeit ist die Darstellung von Oberflächen mit jeweils gleichem Potential (Isocontur)



Aus: A. Leach,  
Molecular  
Modelling,  
2nd ed.



## Molekulares Elektrostatisches Potential (3)

Die Kenntnis dieser Oberflächenladungen ermöglicht ihrerseits die Bestimmung von Atomladungen (z.B. für Kraftfelder)

→ ESP derived atomic charges

Diese müssen wiederum die elektrischen Multipole wiedergeben (iteratives Verfahren)

Literatur zu verschiedenen Ansätzen hierfür:

Cox & Williams *J. Comput. Chem.* **2** (1981) 304

Bieneman & Wiberg *J. Comput. Chem.* **11** (1990) 361

CHELPG Verfahren

Singh & Kollman *J. Comput. Chem.* **5** (1984) 129

RESP Verfahren → Ladungen für das AMBER Kraftfeld

## Molekulares Elektrostatisches Potential (4)

Atomzentrierte Ladungen aus Populationsanalysen der Wellenfunktion weisen eine stärkere Abhängigkeit von dem verwendeten Basissatz auf, als dies bei ESP Rechnungen der Fall ist.

Tendenziell sind die Ladungen auf Wasserstoffatomen betragsmäßig um so größer, je größer der verwendete Basissatz ist.  
→ Elektronendichte „verschmiert“, da nun mehr Basisfunktionen auf den H-Atomen zur Verfügung stehen.

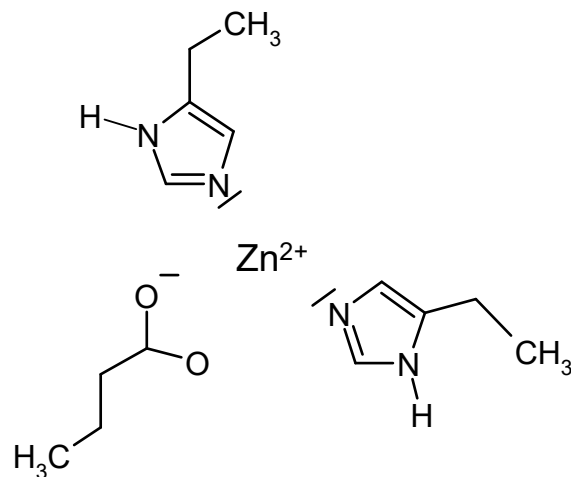
Üblicherweise verwendet man RHF/6-31G\* für ESP Fits.

Atomladung von Kohlenstoff in CH<sub>4</sub>

Basissatz	Anzahl Basisfunktionen	Mullikenanalyse	ESP
STO-3G	9	-0.26	-0.38
3-21G	17	-0.80	-0.45
6-31G(d,p)	35	-0.47	-0.36
6-311++G(2d,2p)	69	-0.18	-0.36
aug-cc-pVDZ	61	+0.63	-0.35

# Molekulares Elektrostatisches Potential (5)

RESP Beispiel des aktiven Zentrums des Angiotensin converting enzyme



	3-21G	6-31G*	6-311++G(2d,2p)	aug-cc-pVDZ
Glu O	-0.993	-0.911	-0.887	-0.876
His N	-0.497	-0.434	-0.385	-0.371
His N	-0.405	-0.364	-0.306	-0.299
Zn	+1.307	+1.239	+1.207	+1.172
Basisfunktionen	255	385	849	776

# Eigenschaften allgemein (1)

Viele molekulare Eigenschaften ergeben sich als Antwort des Moleküls auf eine äußere Störung:

z.B. Entfernen eines Elektrons → Ionisationspotential

Allgemein läßt sich eine Störung durch ein äußeres Feld in einer Taylor Reihe entwickeln. Im Falle eines äußeren elektrischen Feldes  $F$  erhält man ein induziertes Dipolmoment  $\mu_{ind}$ :

$$\mu_{ind} = \mu_o + \alpha F + \frac{1}{2} \beta F^2 + \dots$$

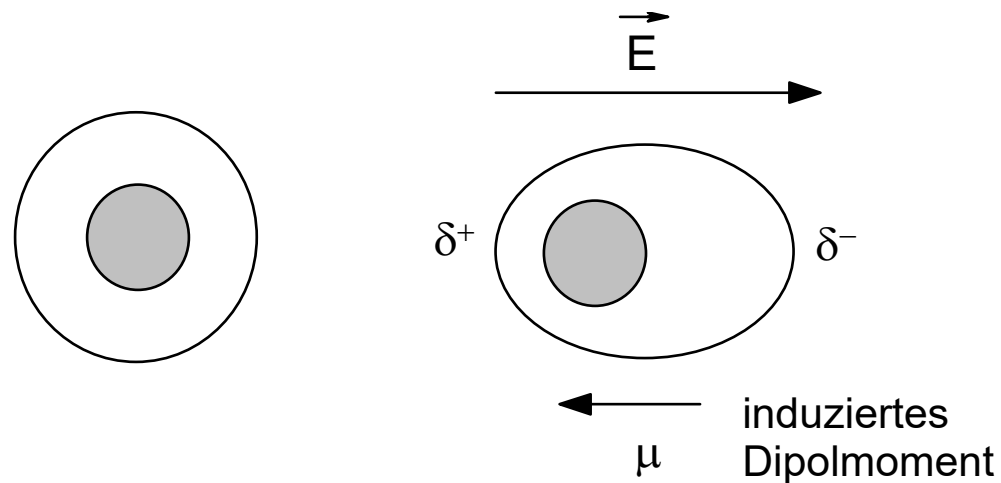
$\mu_o$       permanentes Dipolmoment des Moleküls

$\alpha$         Polarisierbarkeit

$\beta$         (erste) Hyperpolarisierbarkeit

# Hyperpolarisierbarkeit

Durch das elektrische Feld  $E$  des Lichtes werden die Elektronen und Kerne eines Moleküls gegeneinander verschoben, was ein Dipolmoment  $\mu$  induziert. Dieses ist proportional zum elektrischen Feld und der Polarisierbarkeit  $\alpha$ .



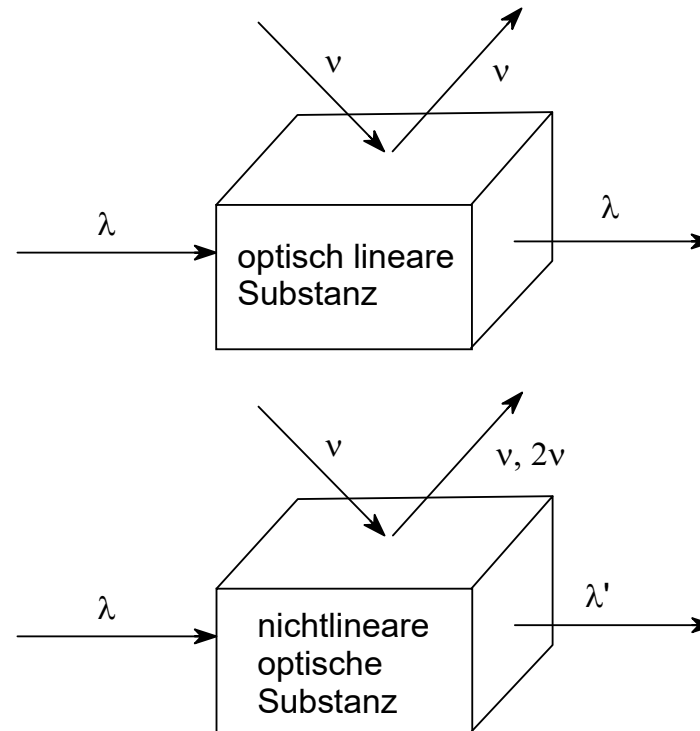
Bei hohen Feldstärken kommt die Hyperpolarisierbarkeit  $\beta$  hinzu.

# Nichtlineare Optik

Experimentelle Beobachtung:

Verbindungen mit hoher Hyperpolarisierbarkeit verändern die Wellenlänge bzw. Frequenz des eingestrahlten Lichtes

Technisch von Bedeutung sind nichtlineare optische Substanzen in der Laser-Anwendung (Durchstimmen von Wellenlängenbereichen) und als optische Schaltelemente (optischer Computer)



## Eigenschaften allgemein (2)

Auswahl von Eigenschaften die sich aus den  $n$ -ten Ableitungen der Energie nach äußeren Feldern berechnen lassen

Elektr.	Magn.	K.Spin	Koord.	Eigenschaft
0	0	0	0	Energie
1	0	0	0	Elektrisches Dipolmoment
0	1	0	0	Magnetisches Dipolmoment
0	0	1	0	Hyperfeinkopplungskonstanten
0	0	0	1	Gradient der Energie (Optimierung)
2	0	0	0	Elektrische Polarisierbarkeit
3	0	0	0	(erste) Hyperpolarisierbarkeit
0	0	0	2	harmonische Schwingungen (IR)
1	0	0	1	IR Absorptionsintensitäten
1	1	0	0	Circularer Dichroismus (CD)
0	0	2	0	Kernspin-Kopplung (J)
0	1	1	0	Kernmagnetische Abschirmung

## Harmonische Schwingungen (IR)

ergeben sich als 2. Ableitung der Energie nach den Koordinaten

Prinzipieller Fehler ist die Vernachlässigung von anharmonischen Schwingungsanteilen. Außerdem hat die Molekülgeometrie (Bindungslängen) Einfluß.

Berechnete IR-Frequenzen sind deshalb (fast) immer zu hoch (ca. 10% gegenüber den experimentellen Wellenzahlen  $\text{cm}^{-1}$ )

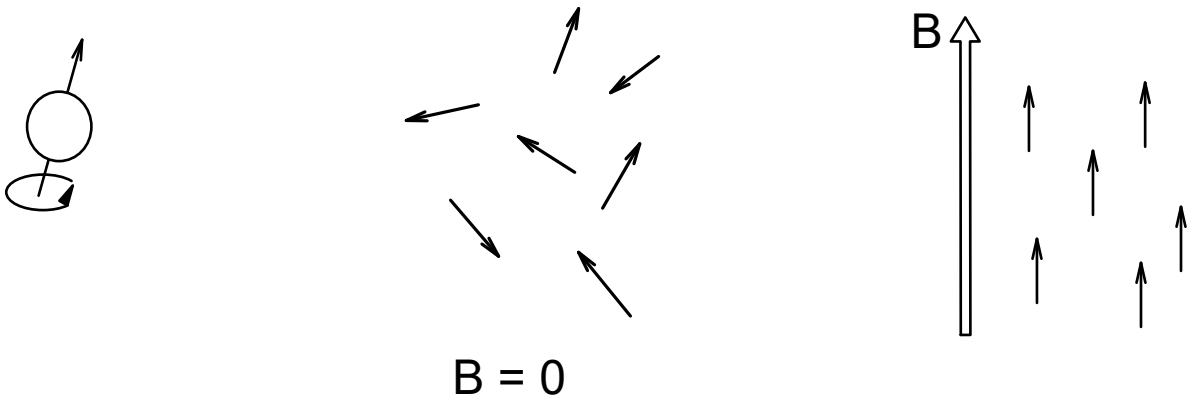
Es gibt deshalb statistisch ermittelte Skalierungsfaktoren für die jeweilige Methode/Basissatz: (RMS Fehler  $\text{cm}^{-1}$ )

HF/3-21G	0.908	87
HF/6-31G*	0.895	50
AM1	0.953	126
PM3	0.976	159
B3LYP/6-31G*	0.961	34
B3PW91/6-31G*	0.957	34



# Kernresonanzspektroskopie NMR (1)

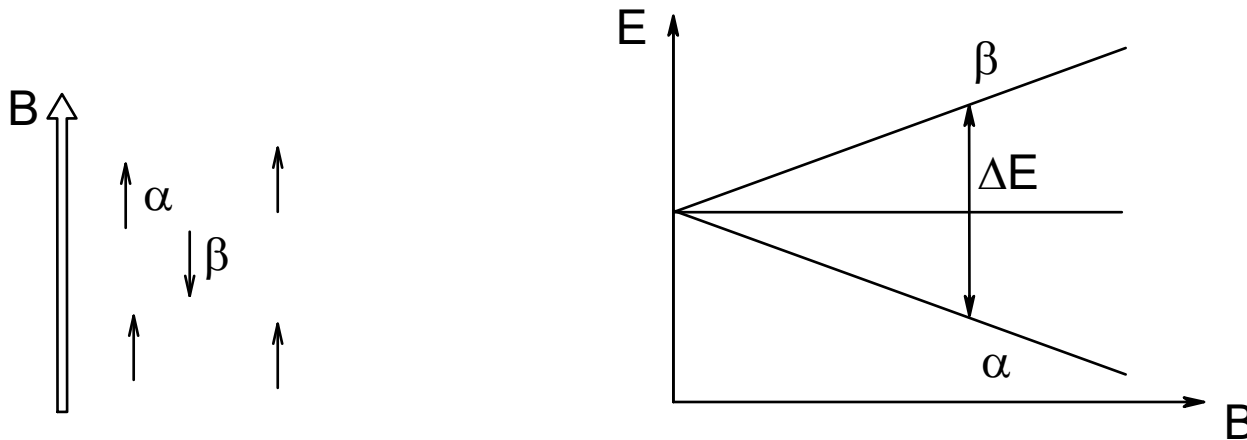
Bestimmte Atomkerne weisen einen sog. Kernspin auf, der die Eigenschaft eines rotierenden Kreisels hat, der wiederum ein magnetisches Moment besitzt. Im feldfreien Raum sind diese Momente zufällig orientiert und nehmen beim Anlegen eines äußeren elektromagnetischen Feldes bestimmte Richtungen ein.



Welche Orientierung des Spins im Magnetfeld die energetisch günstigste ist, hängt von der jeweiligen Spinquantenzahl des jeweiligen Isotops ab.

## Kernresonanzspektroskopie NMR (2)

Für Kerne mit Spin  $\frac{1}{2}$  gibt es zwei Orientierungsmöglichkeiten: mit dem Feld und entgegen dem Feld. Entsprechend ist die Energie von  $\beta$  größer als  $\alpha$ .



Der Energieunterschied für einen Übergang zwischen den Zuständen  $\alpha$  und  $\beta$  hängt außerdem von der Stärke des äußeren Magnetfeldes  $\mathbf{B}$  ab:

$$\Delta E = h\nu = \gamma \frac{h}{2\pi} \mathbf{B}_o \quad \gamma \text{ gyromagnetisches Verhältnis}$$

## Kernresonanzspektroskopie NMR (3)

Der Spin und das gyromagnetische Verhältnis  $\gamma$  sind für jedes Isotop verschieden

Kern	Spin	natürliche Häufigkeit [%]	$\gamma$ [ $10^7$ rad/Ts]	Resonanzfrequenz bei $B_0=2.35$ T [Mhz]	relative Empfindlichkeit
$^1\text{H}$	$\frac{1}{2}$	99.985	26.7519	100.00	1.00
$^2\text{H}$	1	0.015	4.1066	15.351	$9.65 \cdot 10^{-3}$
$^6\text{Li}$	1	7.42	3.9371	14.716	$8.5 \cdot 10^{-3}$
$^{19}\text{F}$	$\frac{1}{2}$	100.0	25.1815	94.077	0.843
$^{13}\text{C}$	$\frac{1}{2}$	1.108	6.7283	25.144	$1.59 \cdot 10^{-2}$
$^{14}\text{N}$	1	99.634	1.9338	7.224	$1.01 \cdot 10^{-3}$
$^{15}\text{N}$	$\frac{1}{2}$	0.366	-0.4835	10.133	$1.01 \cdot 10^{-3}$
$^{17}\text{O}$	$\frac{5}{2}$	0.037	-3.6280	13.557	$2.91 \cdot 10^{-2}$
$^{31}\text{P}$	$\frac{1}{2}$	100.0	10.8394	60.481	$6.63 \cdot 10^{-2}$

## Kernresonanzspektroskopie NMR (4)

Hinzu kommt, daß die Niveaus  $\alpha$  und  $\beta$  gemäß der Boltzmann-Statistik populiert sind:

$$\frac{N_{\beta}}{N_{\alpha}} = e^{-\Delta E/kT}$$

Bsp.  $^1\text{H}$  (Proton) bei  $B_0 = 1.41$  Tesla:  $\Delta E = 2.4 \cdot 10^{-2}$  J/mol  $\rightarrow N_{\beta} = 0.9999904 \cdot N_{\alpha}$   
bei  $B_0 = 7.05$  Tesla:  $\rightarrow N_{\beta} = 0.99995 \cdot N_{\alpha}$

Da im NMR-Experiment Übergänge zwischen den Niveaus beobachtet werden, sollte  $\Delta E$  möglichst groß sein. Dies erfordert möglichst starke Magnete ( $B_0$  groß).

Supraleitende Magnete erreichen heutzutage bis zu 21 Tesla  
Entsprechend wird das Magnetfeld oft auch als Resonanzfrequenz angegeben: 2.3 T entsprechen 100 Mhz, 21.1 T 900Mhz

## Kernresonanzspektroskopie NMR (5)



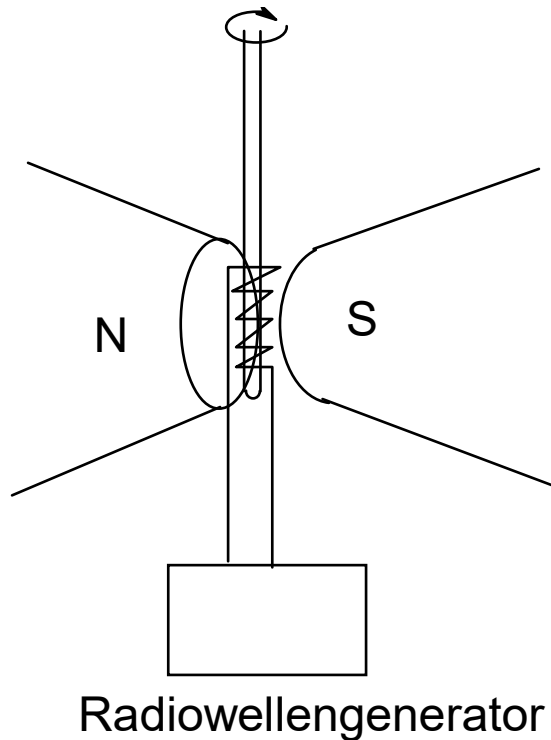
Aktuelle NMR Geräte  
teilweise bis 900 MHz

Kühlung des supraleitenden Magneten  
mit flüssigem Stickstoff (77K) bzw.  
Helium (4K).  
Entsprechend aufwendig und teuer ist  
der Unterhalt.



# Kernresonanzspektroskopie NMR (6)

## Aufbau eines NMR-Spektrometers



Über den Radiowellengenerator läßt sich das äußere Magnetfeld  $B_o$  variieren

$$B_{\text{lokal}} = B_o (1 - \sigma)$$

Die jeweilige Abschirmkonstante  $\sigma$  hängt von der chemischen Umgebung des jeweiligen Atoms ab und bestimmt dessen Resonanzfrequenz  $\nu$

$$\nu = \frac{\gamma}{2\pi} B_o (1 - \sigma)$$

## Kernresonanzspektroskopie NMR (7)

Die Abschirmungskonstante  $\sigma$  berechnet sich (im Idealfall eines einzelnen Protons) nach der Lamb Formel zu

$$\sigma = \frac{\mu_0 e^2}{3m_e} \int_0^\infty r \rho(r) dr$$

in Abhängigkeit von der Elektronendichte  $\rho(r)$ .  $\rightarrow \psi^2$

Damit man Messungen zwischen Geräten unterschiedlicher Magnetfeldstärke  $B_0$  vergleichen kann wird die chemische Verschiebung  $\delta$  als Verhältnis in ppm (parts per million) angegeben:

$$\delta = \frac{\nu_{\text{Substanz}} - \nu_{\text{Standard}}}{\nu_0} \cdot 10^6 \quad \text{mit } \nu_0 = \text{Betriebsfrequenz des Spektrometers}$$

## Kernresonanzspektroskopie NMR (8)

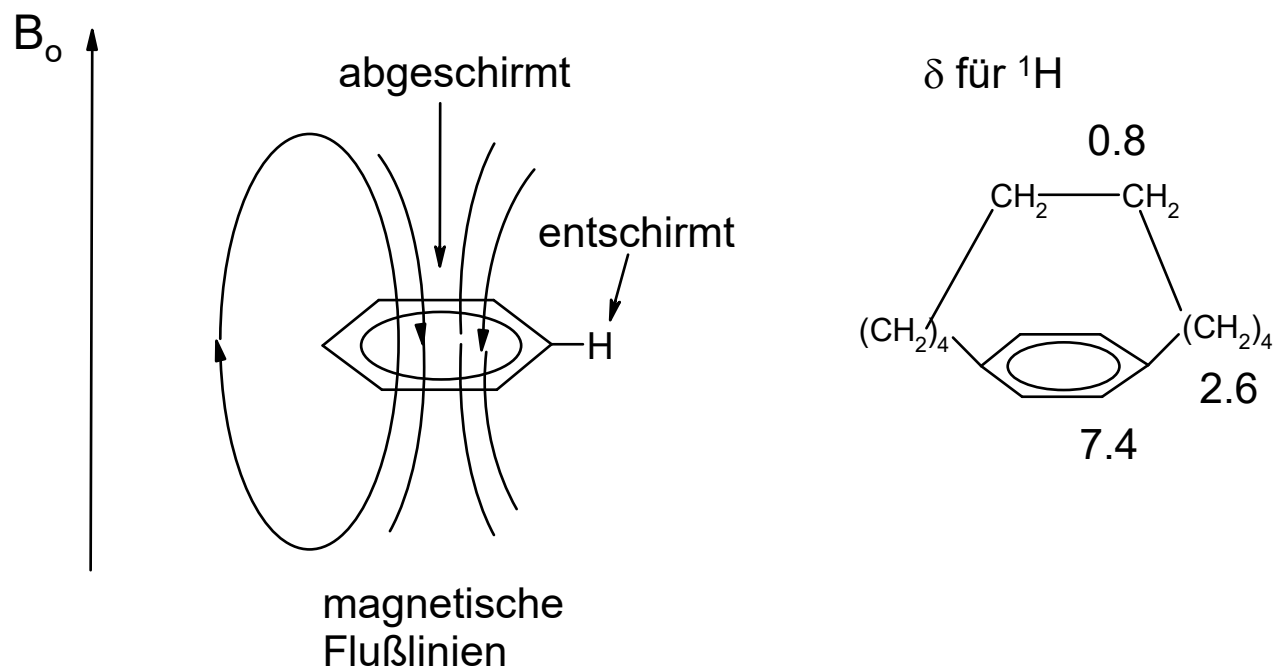
Bei realen Verbindungen setzt sich die Abschirmungskonstante additiv aus verschiedenen Größen zusammen:

- Der diamagnetischen Abschirmung  $\sigma_{\text{dia}}$  bei einer kugelsymmetrischen Ladungsverteilung um den Kern
- Der paramagnetischen Abschirmung  $\sigma_{\text{para}}$  die die nicht-isotrope Ladungsverteilung berücksichtigt. Diese ist proportional zu  $1/\Delta E$ . Zu ihrer Berechnung benötigt man u.a. die Wellenfunktion der angeregten Zustände des Moleküls.
- Die magnetische Anisotropie von Nachbargruppen  $\sigma_{\text{N}}$  die aufgrund von Elektronegativitätsunterschieden zustande kommt.
- Dem Ringstromeffekt von Aromaten  $\sigma_{\text{R}}$
- Dem elektrischen Effekt  $\sigma_{\text{e}}$  durch geladene Gruppen ( $-\text{NH}_3^+$ ,  $\text{COO}^-$ )
- Den intermolekularen Wechselwirkungen  $\sigma_{\text{i}}$  durch Wasserstoffbrücken und das Lösungsmittel (Dielektrikum).



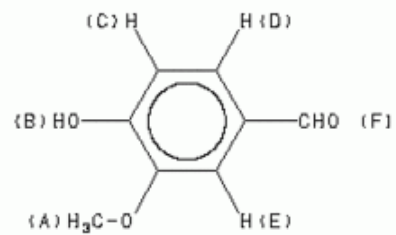
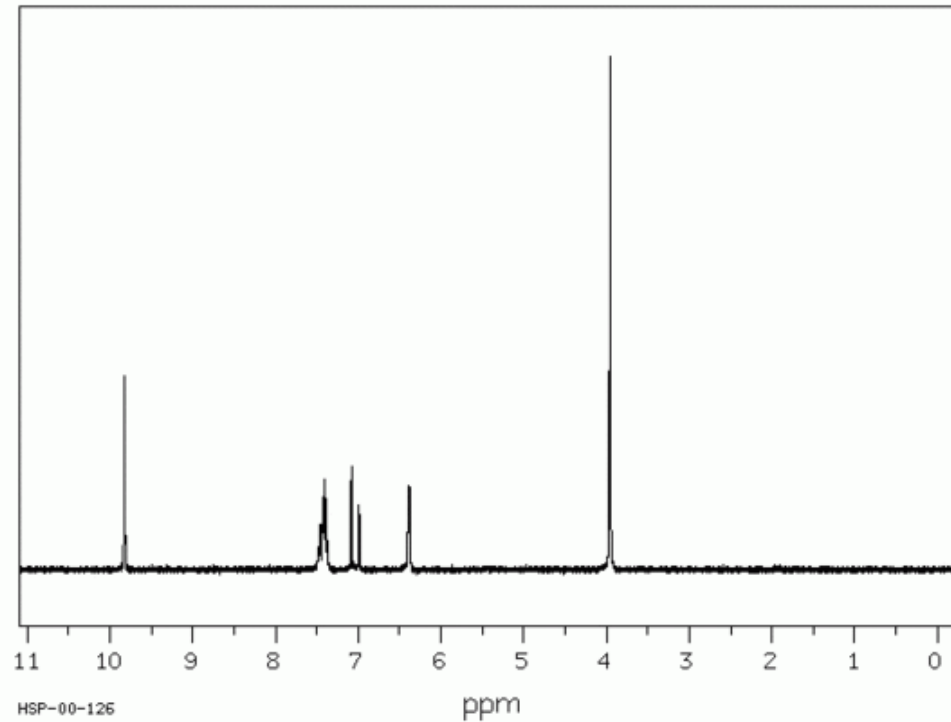
# Kernresonanzspektroskopie NMR (9)

## Ringstromeffekt von aromatischen Systemen



# Kernresonanzspektroskopie NMR (10)

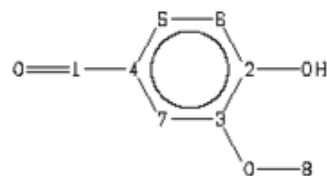
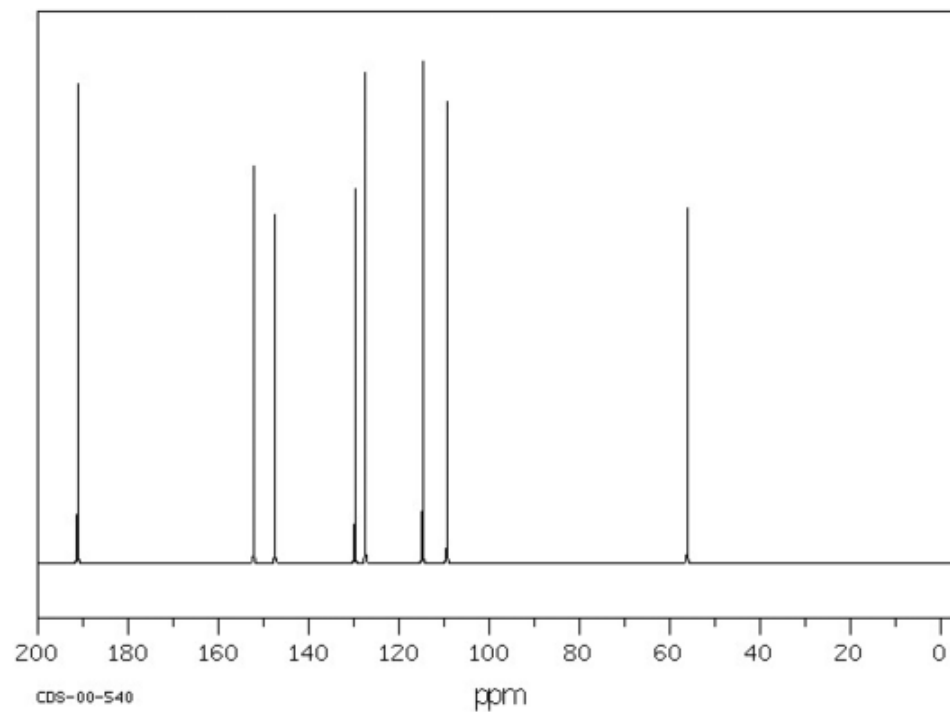
## $^1\text{H}$ -NMR Spektrum von Vanillin



Location	Shift (ppm)
A	3.959
B	6.39
C	7.047
D	7.42
E	7.42
F	9.823

# Kernresonanzspektroskopie NMR (11)

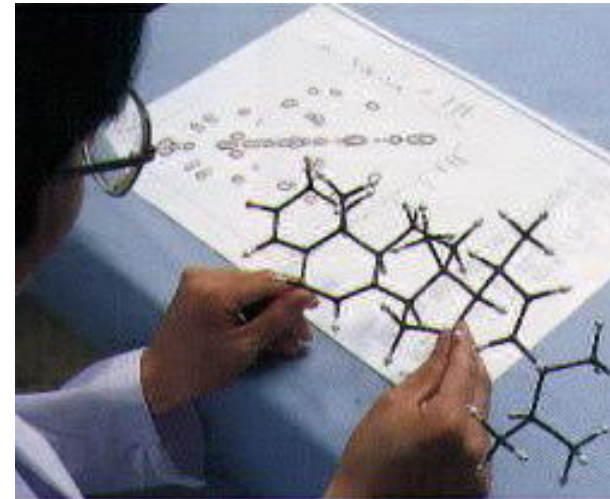
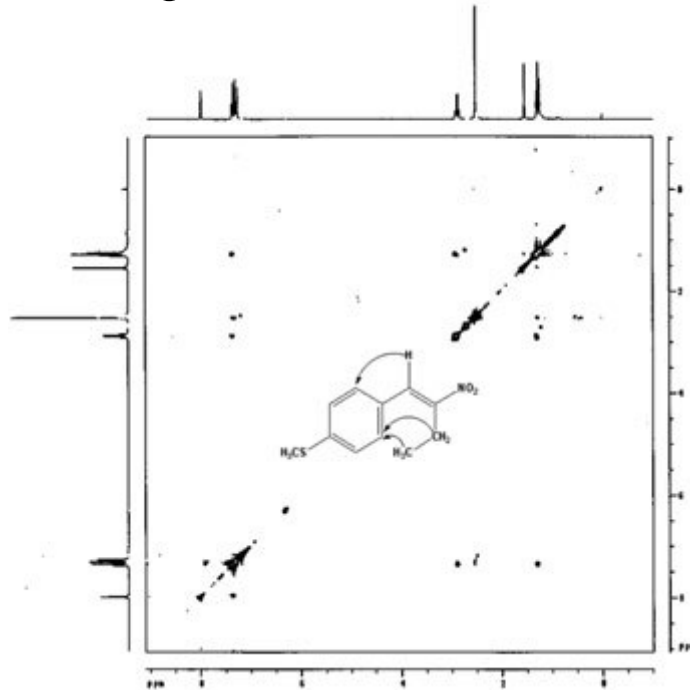
## $^{13}\text{C}$ -NMR Spektrum von Vanillin



ppm	Int.	Location
191.21	955	1
152.18	791	2
147.50	692	3
129.77	746	4
127.49	975	5
114.75	1000	6 *
109.34	920	7 *
56.10	706	8

# Kernresonanzspektroskopie NMR (12)

Aus der Intensität der *cross-peaks* von 2D-NMR Spektren läßt sich die räumliche Nachbarschaft von Kernen ableiten. Die Signalintensität fällt mit  $1/r^6$  ab.

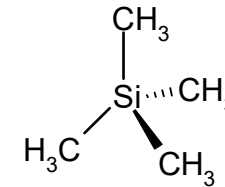


Ermöglicht Strukturaufklärung von kleinen Proteinen  
(bis etwa 100 Aminosäuren)

# Kernresonanzspektroskopie NMR (13)

Verschiebungsbereiche von Isotopenkernen

Kern	$\delta$ typisch [ppm]	Referenzstandard
$^1\text{H}$	0 - 10	TMS (Tetramethylsilan)
$^{13}\text{C}$	0 - 200	TMS (Tetramethylsilan)
$^{14}\text{N}$	900 - 1000	Nitromethan oder $\underline{\text{N}}\text{H}_4\text{NO}_3$ , beide extern
$^{31}\text{P}$	-100 - +250	$\text{H}_3\text{PO}_4$ , extern



Da die 12 H-Atome von TMS chemisch alle gleich sind, wird im NMR Spektrum ein starkes Singulett Signal beobachtet.

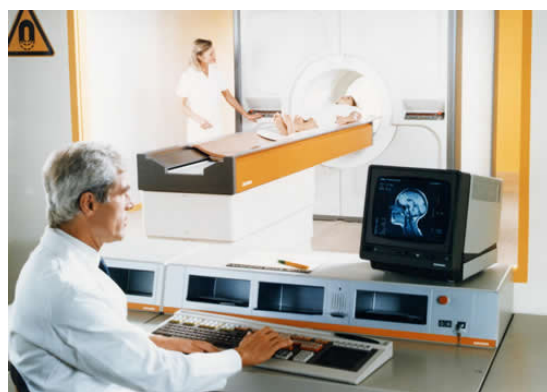
Praktische Umsetzung der Berechnungen von NMR Verschiebung v.a. durch W. Kuzzelnigg und P. Schleyer

# Kernresonanzspektroskopie NMR (14)

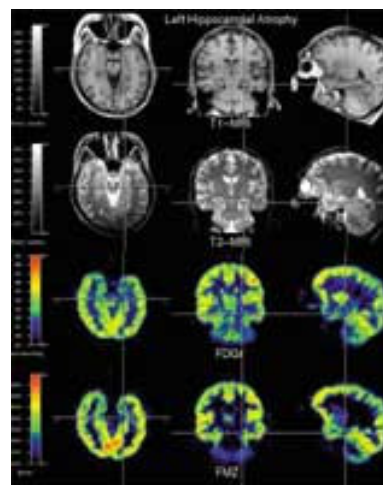
Während für die analytische Zwecke in der präparativen organischen Chemie v.a. die  $^1\text{H}$ ,  $^2\text{H}$  und  $^{13}\text{C}$  Verschiebungen von Bedeutung sind, werden für die Strukturaufklärung von Proteinen zweidimensionale NMR Spektren benötigt.

Bsp. Kopplung zwischen Kernen ist abstandsabhängig und gibt deshalb Information über die Entfernung.

In der Medizin wiederum ist die räumliche Verteilung von Bedeutung.



Werksfoto Siemens (um 1985)



# Vorhersage komplexerer Moleküleigenschaften

Komplexe Eigenschaften lassen sich indirekt bei Kenntnis der experimentelle Werte für einen Datensatz von Molekülen durch wissensbasierte Methoden bestimmen

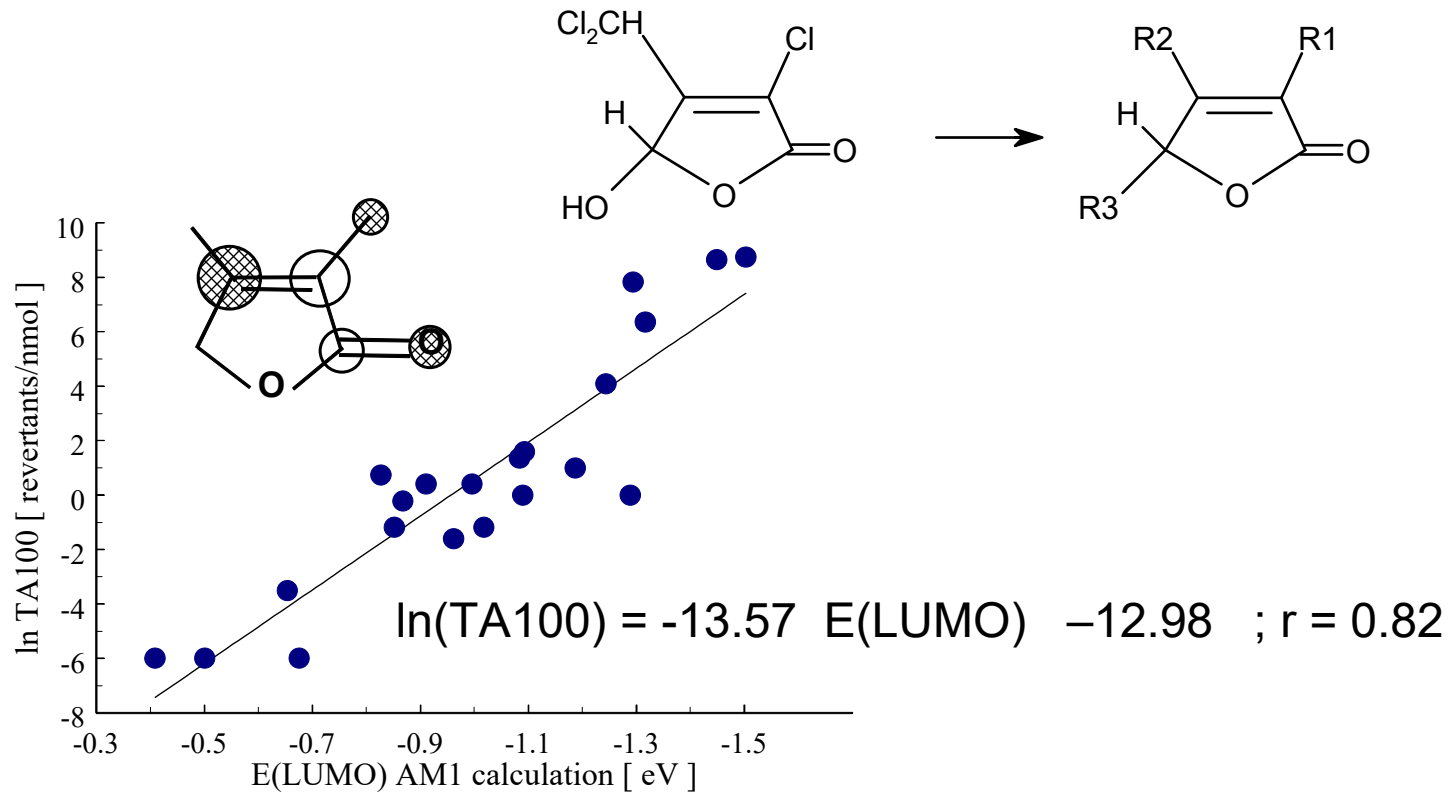
Allgemeiner Ansatz:

1. Statistische Ermittlung signifikanter Variablen
2. Aufstellen einer Regressionanalyse und Fit der Koeffizienten an einem hinreichend großen Satz von Daten  
quantitative structure activity relationship (QSAR)  
quantitative structure property relationship (QSPR)
3. Validierung an einem Testsatz von Daten die nicht für den Fit benutzt wurden, oder *cross-validation*

# Quantum QSAR

Generierung der molekularen Eigenschaften für die QSAR-Gleichung aus quantenmechanischen Daten.

Bsp: Mutagenizität von MX-Verbindungen

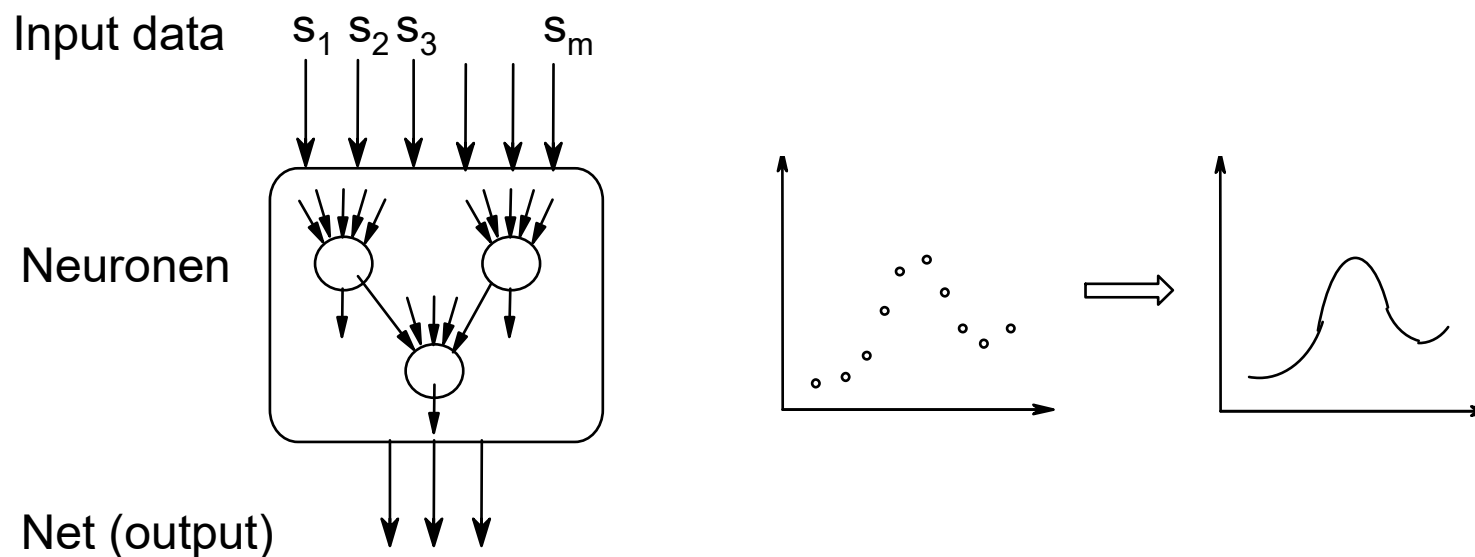


Lit.: K. Tuppurainen et al. *Mutat. Res.* **247** (1991) 97.



# Neuronale Netze (1)

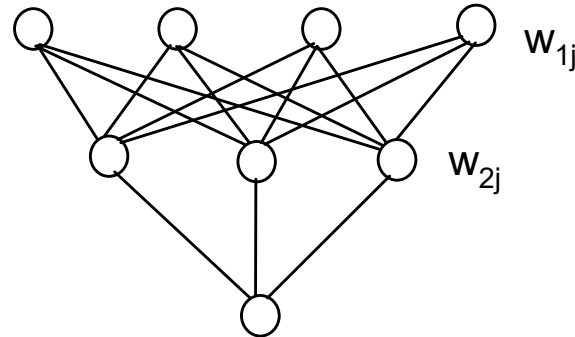
Neuronale Netze sind eine gängige Implementierung der “künstlichen Intelligenz”. Der Name leitet sich aus der netzwerkähnlichen Verknüpfung zwischen den Schaltelementen (Neuronen) innerhalb des Systems ab.



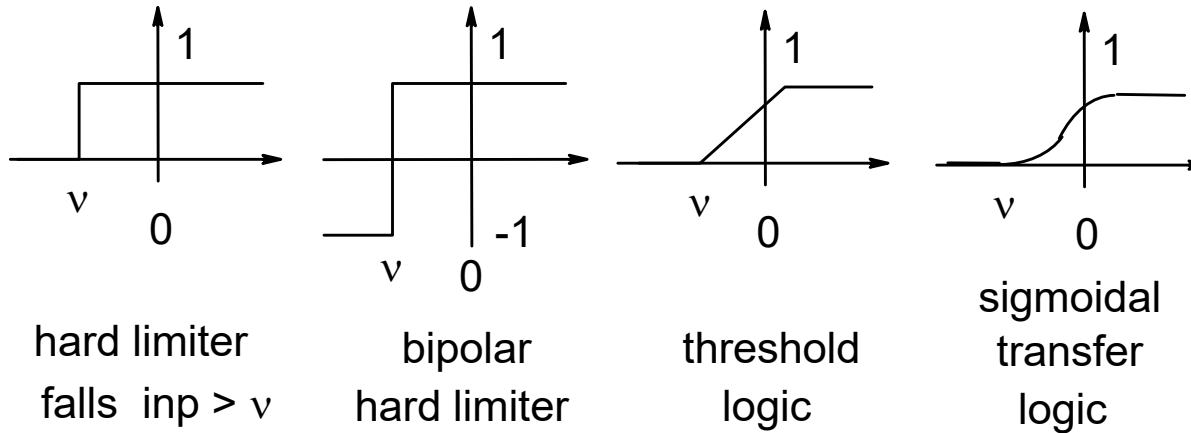
Gegenüber Regressionsgleichungen lassen sich mit neuronalen Netzen auch inter-korrelierte Deskriptoren verarbeiten.

# Neuronale Netze (2)

Ein (konventionelles) neuronales Netz besteht aus einem *input layer*, ein bis mehreren *hidden layers*, sowie dem *output layer*



Außerdem kann die Art der Signalweitergabe zwischen den Neuronen unterschiedlich sein:



# LogP aus einem neuronalen Netz (1)

LogP: Wasser/n-Octanol Verteilungskoeffizient

$$-4 < \log P < 8$$

hydrophil

lipophil

Idee: Der Verteilungskoeffizient ist durch die elektronischen Eigenschaften eines Moleküls bestimmt die quantenchemisch berechnet werden können.

Zu den Eigenschaften gehören: Dipolmoment, mittlere Polarisierbarkeit, Moleküloberfläche, Molekülvolumen, Globularität, Anzahl und Typ von Heteroatomen (N, O, P, S, F, Cl, Br, I), minimales und maximales elektrostatisches Potential und statistische Größen davon (Mittelwerte und Varianz).

## LogP aus einem neuronalen Netz (2)

*backpropagation net* mit 16 Input Neuronen, 25 Neuronen im *hidden layer* und 1 Output Neuron das den logP angibt.

Als Trainings Set wurden 980 zufällig ausgewählte Moleküle aus einer Datenbank von 1085 Verbindungen gewählt. Die restlichen 105 dienten als Test Set.

Eigenschaften berechnet mit	AM1	PM3
Trainings Set	$r^2=0.965$	$r^2=0.883$
Standartabweichung	0.41	0.45
Test Set	$r^2=0.902$	$r^2=0.830$
Standartabweichung	0.53	0.67

Lit: Breindl, Beck, Clark, Glen, *J.Mol. Model.* **3** (1997) 142

# Moleküleigenschaften Zusammenfassung

Alle Moleküleigenschaften die sich auch aus spektroskopischen Messungen bestimmen lassen sind direkt aus der Wellenfunktion berechenbar.

→ Wechselwirkungen zwischen elektromagnetischer Strahlung und der Materie (= Kerne + Elektronen)

Sonstige Moleküleigenschaften (z.B. Bindungsaffinitäten, logP, usw.) kann man mit Hilfe von Regressionsgleichungen, neuronalen Netzen, oder anderen Methoden des „statistical learnings“ anhand experimentell ermittelter Werte von (möglichst) vielen Molekülen kalibrieren und dann für neue Verbindungen vorhersagen.

→ *in silico* screening