The Influence of Conformational and Associative Effects on the QSPR Descriptors of Oligopeptide Derivatives

Habilitationsschrift

zur Erlangung des akademischen Grades
doctor rerum naturalium habilitatus (Dr. rer. nat. habil.)

vorgelegt der

Mathematisch-Naturwissenschaftlich-Technischen Fakultät
(mathematisch-naturwissenschaftlicher Bereich)
der Martin-Luther-Universität Halle-Wittenberg

von

Dr. rer. nat. MONIKA PLASS
geb. am 3. Juni 1965 in Potsdam

Gutachter:
1. Prof. A. Kolbe, Martin-Luther-Universität
2. Prof. G. Oehme, Universität Rostock
3. Prof. M. H. Abraham, University College London

Halle (Saale), den 13. Dezember 2000

CONTENTS

ABBREVIATION

Substances and residues

The amino acid and peptide derivatives were abbreviated according to the recommendation of the IUPAC[1] using the three letter code. If not particularly indicated the amino acids were L- configurated. An unspecified amino acid is abbreviated with Xaa.

Ac	Acetyl
AcN	Acetonitrile
Boc	tert. Butyloxycarbonyl
DHPC	1.2-dihexadecyl-sn-glycero-3-phosphoryl choline
DMPC	1.2-dimyristoyl-sn-glycero-3-phosphoryl choline
DPPC	1.2-dipalmitoyl sn-glycero-3-phospho choline
MeOH	Methanol
NMA	N-methyl acetamide
HFBA	Heptafluorobutyric acid
HFiP	1.1.1.3.3.3-hexafluoropropanol
OBzl	Benzoxy
OEt	Ethoxy
OMe	Methoxy
OnBu	n-Butyl ester
TFA	Trifluoroacetic acid
TFE	1.1.1-trifluoroethanol
Tos	Tosyl
Z	Benzyloxycarbonyl

General abbreviation

AM1	Semi-empirical method
c	Concentration in mol/l
CD	Circular Dichroism
eq.	Equation
E_r	Relative energy of a molecule, in kJ/mol
Fig.	Figure
FTIR	Fourier Transform Infrared
HPLC	High Performance Liquid Chromatography
M_M	Molecular Mass, in g/mol
NMR	Nuclear Magnetic Resonance
QSPR	Quantitative Structure-Property Relationship
r^2	Regression coefficient
S. D.	Standard deviation
T	Temperature, either given in °C or K
UV	Ultraviolett spectroscopy
VCD	Vibrational Circular Dichroism

QSPR

BBB	Blood-brain barrier
c, r, s, a, b, v	Coefficients characterising the specific partitioning system in eq. 4
CLOGP	Calculated Logarithm of the Partitioning coefficient
LFER	Linear Free Energy Relationship
logP	Logarithm of a partitioning coefficient

$\log P_{Oct}$	Logarithm of the partitioning coefficient in the system octan-1-ol/water
logSP	Logarithm of any solute related property
π_2^H	Dipolarity/polarisability in eq. 4
R_2	Excess molar refraction in eq. 4
$\sum\alpha_2^H$	Effective hydrogen bond acidity in eq. 4
$\sum\beta_2^H$	Effective hydrogen bond basicity in eq. 4
UCL database	Database of solute descriptors at the University College London
V_x	Characteristic McGowan volume, given in l/100 mol in eq. 4

HPLC measurements

CD	Permethylated β-cyclodextrine column (see: Table E2)
CHI	Chromatographic Hydrophobicity Index from gradient HPLC measurements
CN	Cyanol column (see: Table E2)
IAM	Immobilised artificial membrane column (Table E2)
In	ODS2 column, Inertsil (see: Table E2)
φ_0	Chromatographic Hydrophobicity Index from isocratic HPLC measurements
logk	Logarithm of the retention factor k of a compound
logk', $\log k_0$	Logarithm of the retention factor, intercept at 100% water received from linear correlation of logk vs. %organic modifier (isocratic HPLC measurements)
$\log k_W$	Logarithm of the retention factor at 100% water
ODS	Octadecyl silicate
Pro	ODS2 column, Protegy (see: Table E2)
RP	Reversed phase
S	Slope value, received from linear correlation of logk vs. %organic modifier (isocratic HPLC measurements)
t_G	Gradient time; time interval of linear increase of the organic modifier from 0 to 100 %
t_R	Retention time of a compound from the gradient HPLC measurements
w	Bandwidth of the chromatographic signal

Spectroscopic measurements

C_{10}	Intramolecularly hydrogen bonded ten membered associate ring
C_5	Intramolecularly hydrogen bonded five membered associate ring
C_7	Intramolecularly hydrogen bonded seven membered associate ring
c_A	Equilibrium concentration of the proton donor
c_{A0}	Initial concentration of a proton donor compound
c_{ass}	Equilibrium concentration of the associated species
c_B	Equilibrium concentration of a proton acceptor
c_{B0}	Initial concentration of a proton acceptor
d	Thickness of the IR cell, in mm
δ	Proton chemical shift, in ppm
Δδ	Relative proton chemical shift, in ppm
Δν	Relative shift of an IR signal, in cm^{-1}
Δν/ΔT	Relative shift of an IR signal with the change of the temperature, in cm^{-1}/K
ε	Dielectric constant
ε_{Mol}	Molar extinction coefficient (integral or absolute), in l/mol*mm
K	Equilibrium constant
$K_{1:1}$	Equilibrium constant for a 1:1 associate formation
$K_{1:2}$	Equilibrium constant of 1:2 associates

K_{Ac}	Equilibrium constant for the 1:1 association on the acetamide function
K_{app}	Apparent equilibrium constant
K_{Ester}	Equilibrium constant for the 1:1 association on the ester function
K_{OH}	Equilibrium constant using an alcohol as proton donor, calculated from the decrease of the intensity of the IR signal
$K_{Peptide}$	Equilibrium constant for the 1:1 association on the peptide function
K_{all}	Equilibrium constants for the 1:1 association on all acceptor functions
K_Z	Equilibrium constant for the 1:1 association on the urethane function
$\log K_{1:1}$	Logarithm of the equilibrium constant for 1:1 association
MFP	Magnetic Field Perturbation
T_m	Main phase transition temperature, in °C
ν	Wavenumber, in cm^{-1}
v	Diffusion rate

A INTRODUCTION

A.1 Peptide drugs

The peptide comprises a great variety of biologically active linear and cyclic biopolymers with diverse functions which can be divided according to their functions into different classes such as antibiotics, enzyme inhibitors and substrates, hormons and regulatory factors, peptide alkaloids, toxins and sweeteners[2]. Most of those bioactive peptides consist of a larger number of amino acid residues but there are also a small drugs consisting of up to three amino acid residues. Recently, it was concluded from animal studies that some tripeptides, such as the thyrotropin-releasing hormon (TRH) and the C-terminal fragment of oxytocin, prolyl-leucyl-glycinamide (PLG), show antidepressant effects[3]. Principal arguments against the application of peptide drugs were based on the susceptibility of these compounds to enzymatic attack in the gut and the circulation with a resulting short half-life as well as the need to administer them by the route of injection. Furthermore, the drugs have to pass rapidly through cell membranes and tissues without being cleared from the system. Long hydrophilic amino acid chains block the transport across the biological membranes: Absorption does not correlate with bioactivity. Crossing membranes may provoke denaturation or degradation and may even result in potential immunogenicity or even toxicity[4]. Hence the transport of the peptide through the membrane plays a key position in its applicability. In the case of neuropeptides in particular the blood-brain barrier (BBB) as a special membrane seems considerable.

A.1.1 Transport of peptide drugs through cell membranes

There are two distinctly different pathways for polypeptides to cross the epithelium: the transcellular or the paracellular route[5]. The paracellular route involves diffusion of the solute in the extracellular space between adjacent cells restricted by the presence of tight junctions. Transport by this route is dependent on the size and charge of the molecules and contributes little to absorption of other than small solutes.

The transcellular mechanism involves the movement of the solute across the apical cell membrane -through the cell interior and across the basolateral membrane by either active or passive processes. The active transport requires a carrier or receptor molecule and is thus fairly substrate specific. The choice of the carrier system required to transport a peptide drug through the membrane will be significantly determined by the acidic, basic or amphoteric properties of the drug[6, 7]. Examples for an receptor-mediated transport are peptides like

insulin[8], leucine encephalin[9] and arginine vasopressin[10]. The extraction of the tripeptide thyrotropin-releasing hormon (TRH) seems to work by passive diffusion[11, 12].

In a first approximation those diffusion processes can be modelled successfully by partitioning in octanol like phases[13, 14]. More extensive permeability studies on Caco-2 cell monolayers as a model for intestinal mucosa show that the transport of peptides is rather complex and cannot simply be described by partitioning coefficients in the traditionally used water-octanol system. But the hydrogen bonding potential may be the major determinant of transport across cell membranes. This apparent inconsistency is reasoned by the heterogeneous structure of the cell membrane consisting of an amorphous lipid matrix in which and upon which a variety of membrane proteins are distributed[15]. Thus, if passing the cell membrane the solute will experience different microenvironments. If only referred to interaction with the phospholipids the peptide molecule will first interact with the hydrated polar head groups of the phospholipids which point into the aqueous exterior of the cell. Then it has to pass the apolar membrane interior which demands apolar/hydrophobic properties of the solute. Recently, Burton et al.[16] concluded that a peptide must possess a delicate balance of affinity of the aqueous-membrane interface and a reasonably low desolvation energy in order for it to efficiently cross an epithelial cell membrane.

A.2 Lipophilicity and drug design

Transport and distribution processes within biological systems are to a large extend controlled by lipophilicity. The hydrophobic interior of biological membranes selects the compounds which are able to cross the barrier. Hence hydrophobic and hydrophilic properties of the compounds should be well proportioned to guarantee the effective transport of drugs within biological tissues. The lipophilicity of a compound is dependent on its structural composition and the capability of the structure to interact with its environment. It encodes most relevant intermolecular forces that can take place between a solute and a solvent which are ion-ion and ion-dipole interactions, charge transfer interactions, hydrogen bonds, hydrophobic bonds or Van der Waals interactions.

However, also specific intramolecular interactions do influence the lipophilicity if not just by deactivating parts of the molecule. As a consequence certain conformers might be favoured and statistic conformational distribution is not relevant.

The IUPAC suggested the following definition of the lipophilicity[17, 18]:

„*Lipophilicity* represents the affinity of a molecule or a moiety for a lipophilistic environment. It is commonly measured by its distribution behaviour in a biphasic system, either liquid-liquid or solid liquid systems.“

The most widely used measure for the hydrophobic/lipophilic[19] properties of a compound in many biological partitioning processes[20] is the logarithm of the equilibrium concentration in octanol-water $\log P_{Oct}$ which has been established as a quantitative parameter for the lipophilicity of a compound[14, 21,] Octan-1-ol combines the lipophilic properties of the alkyl chain and the hydrophilic hydroxyl group which is capable to form hydrogen bonds in its molecular structure. These structural features makes it a proper model to simulate the properties of phospholipids which are the major components in the molecular composition of a lipid membrane[22, 23]. In recent times for certain purposes also other solvent systems such as ethylene glycol/water were used for the description of particular partitioning behaviour.

In this regard also the term hydrophobicity is often used interchangeable. The IUPAC made a clear distinction in both definitions:

„*Hydrophobicity* is the association of nonpolar groups or molecules in an aqueous environment which arises from the tendency of water to exclude nonpolar molecules.“

However, the definitions are discussed controversial and are still being far from completeness. Hence it is often found in scientific research papers that both definitions are not pragmatically differentiated. Whereas medical chemists use lipophilicity in order to describe transport processes of compounds in biological systems biophysicists working with spectroscopic methods or investigating interaction properties on molecular surfaces tend to use the term hydrophobicity. This work delivers a contribution to the biophysical aspects and is based on spectroscopic investigations and HPLC measurements. The author will use the term hydrophobicity in order to be uniform with cited papers but being fully aware of the differences in the present definition.

A.3 QSPR analysis -methods

In the approach of the effective drug design Linear Free Energy relationships (LFER) play an essential role since the understanding of the consequences the structural composition on the activity of the drugs allows the prediction of drug efficiency. The basic concept of these empirical approaches is often based on the comparison of the behaviour of structurally similar substances in defined environments or reactions. LFER equations were applied to relate spectroscopic parameters e.g. NMR chemical shift, vibrational frequencies and also solvent

effects with structural features of the compounds with the intention to develop a system of descriptors which helps understanding reaction mechanism and physical properties.

One of the key parameters for delivery of drugs is their transport through human tissues and their partitioning behaviour in the complexity of the chemical compounds. The lipophilic properties of the prodrug decide about their relative affinity of the dissolved compound in an aqueous media and an organic phase immiscible with water represented by logP.

There are different approaches how to analyse logP values and to receive the structural descriptors. The key question is how divide logP into additive fragments correlated to structural features of the molecule.

This can be done by breaking a molecule into atom fragments (CLOGP method)[24, 25] or by using quantum mechanic calculation as a tool to describe the charge density and the induced polarity of a molecule based on the molecular orbitals which makes the approach independent of any experiment[26].

Furthermore, the analysis of experimentally determined logP for a series of compounds can result in logP values for substituents if parent molecules are used as reference moiety or also by dividing the molecules into fragments which deliver a contribution to logP.

Similarly to the Hammett equation[27] the substituent method bases on the correlation of partitioning coefficient P_X of compound X with that of a reference substance. According to eq. 1 the difference of the log P values gives a lipophilicity descriptor π_X which characterises the influence of the substituent X.

$$\log P_X = \pi_X + \log P_H \qquad [1]$$

The fragmental method by Nys and Rekker[28] adds the logP values of molecular subunits F_i. As it was shown by Abraham and Leo[29] this method fails for peptide compounds because the conformational flexibility cannot be recognised. The implication of correction terms f_i improved the success of the method.

$$\log P = \sum_{i=1}^{i=m} f_i + \sum_{i=1}^{i=n} F_i \qquad [2]$$

Fragmentation based on the interaction properties displays logP values on a different level. Here the molecule is not divided into molecular subunits but into interacting subunits. This

approach traces back to the contribution of the major interactions to logP as lipophilicity parameter.

Dunn and Wold[30] have analysed the solvent parameters of solutes by the principal compound analysis. According to their investigations logP is mainly determined by two effects. Firstly, the molecular volume or the surface effect of the solute seems to be important. Secondly, the strength of the polar interaction between the solute and solvent influences the logP value. With regards to peptide and protein investigations Janin[31] used the linear correlation of the free energy of the solvent transfer of an organic solvent into the aqueous phase and the accessible protein surface for the creation of a hydrophobicity scale of amino acids.

The cavity model is the basis for the calculation of logP from molecular properties. The observations of Cramer[32] show that the energy for the formation of a solvent cavity is compensated by the attractive forces between solute and solvent. Thus, logP can be calculated from a set of descriptors which characterise the attractive and repulsive properties of a solute[33, 34, 35].

$$\log P = a\,V + b\pi^* + c\,\beta_H + d\,\alpha_H + e \qquad [3]$$

The concept of descriptors for interacting forces was modified for solution properties and extended by a refractivity term characterising the shape of the molecules[36]. Hence the solvation equation by Abraham eq. 4 covers the influence of the volume with the characteristic McGowan volume V_x, given [in l/100 mol] and the shape of the molecule with the excess molar refraction R_2. The most important solute-solvent interactions for neutral molecules are hydrogen bonds, where the solute may act as a donor/acceptor, and dipole interactions[37]. Dipolar interactions are represented by the dipolarity/polarisability term π_2^H. The hydrogen bonding properties are represented by the effective hydrogen bonding acidity $\Sigma\alpha_2^H$ and basicity $\Sigma\beta_2^H$.

$$\log SP = c + rR_2 + s\pi_2^H + a\Sigma\alpha_2^H + b\Sigma\beta_2^H + \nu V_x \qquad [4]$$

The coefficients c, r, s, a, b and ν characterise the specific partitioning system and are usually determined in preliminary investigations with substances of known descriptors. LogSP can be any solute related property. The importance of this equation lies in its numerous

applications on different partitioning systems. Hence equations for the characterisation of biological systems as well as physicochemical partitioning systems are known. A selection of partitioning systems with known coefficients in the solvation eq. 4 are given in table 1.

Table 1: Ratios of the coefficients for selected partitioning systems

System	r/v	s/v	a/v	b/v
Hexadecane/water[38]	0.15	-0.36	-0.81	1.10
Octanol/water[38]	0.15	-0.28	0.01	-0.91
Blood/brain[39]	0.19	-0.69	-0.72	-1.28
CH_2Cl_2/water[38]	0.0002	0.0051	-0.760	-0.971
skin/ water[40]	0.00	-0.327	-0.354	-1.949
SDS micelles/water[41]	0.194	-0.143	-0.0476	-0.565

Hence, if the partitioning properties of a relevant biological system expressed by the five coefficients in eq. 4 are known the solute property of any compound can be predicted. Conclusively, *in vivo* measurements as in the blood-brain barrier which are usually difficult can be avoided or can be performed only on selected prodrugs. Furthermore, the descriptors are not just empirical values since they are based on interacting properties and they relate to certain spectroscopic properties which also represent solute-solvent interactions. The method offers an enormous potential for drug development since the determination of the solute descriptors can be done in partitioning systems which are easy to handle and to automate.

However, all LFER equations are based on the principle of additivity of partitioning properties when the structure of the molecules are similar. This concept fails if structural peculiarities occur in the partitioning system. Such deviations can be caused by changes of the secondary conformation or in the case of specific intramolecular or intermolecular interactions of the solute and the solvent.

The following example will illustrate that the explanation of the diversion between predicted and experimental logP values is not a trivial task and may demand extensive research combining different methods. Hence the logP data of the herpes drug, acyclovir[42, 43] and its derivatives were investigated. In Fig. 1 the experimentally determined $logP_{Oct}$ values of ACV acyclovir (9-(2-hydroxyethoxymethyl)guanine) derivatives were plotted vs. the calculated $logP_{Oct}$ values based on the Rekker method.

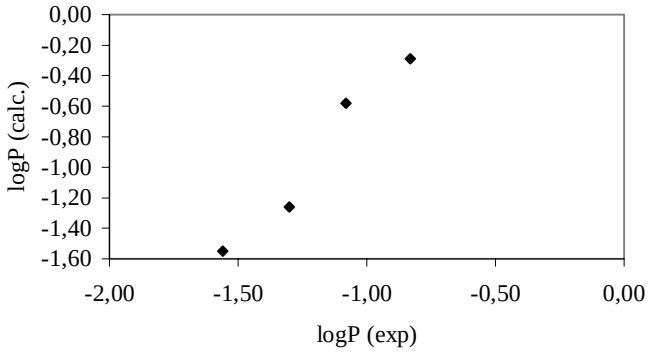


Fig. 1: $LogP_{Oct}$ values determined experimentally and calculated by Rekker's method for acyclovir derivatives[44]

As can be seen easily the results of the $logP_{Oct}$ calculation give no satisfying results. Hence, the question arises why these data cannot be predicted from the fragmental data.

It is known from the literature[45, 46, 47, 48] that guanine derivatives exist in a tautomeric equilibrium as it is shown in Fig. 2.

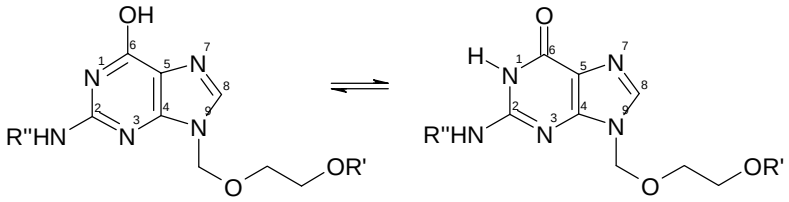


Fig. 2: Tautomeric equilibrium of acyclovir derivatives (ACV: R', R''=H; NAcACV: R'=H, R''=acetyl; diAcACV: R', R''=acetyl)

From quantum mechanical calculations based on different basic sets it is known that guanine derivatives can exist in both tautomers[49, 50, 51]. Our own calculations using the semiempirical AM1 method favours the oxo-amino tautomer for all acyclovir derivatives in the back-folded arrangement of the side chain on N-9 which was identified as the most relevant arrangement of the side chain[44].

Table 2: Energetic differences of the acyclovir derivatives based on AM1 calculations

Substance	ΔE(keto-enol)	Dipole moment in Debye	
	[in kJ/mol]	keto	enol
ACV	-20.8	5.92	2.34
NAcACV	-42.5	8.73	7.61
diAcACV	-45.7	7.90	7.57
DCV		2.24	

In comparison to the calculations performed on 9-methyl guanine with a 6-31 G basis set at the MBPT2 level with zero energy the stabilisation energy in table 2 seems rather large[51]. However, this is a general derivation caused by the AM1 parameterisation which is obviously overestimating the energetic differences between the tautomers.

It is worthwhile mentioning that the thermodynamic stabilisation of the oxo-amido tautomers in the N- acetyl derivatives is caused by an intramolecular hydrogen bond between the N-1-H function and the amide group on C-2 purine ring. The atom distance O...H-1-N is equal to 2.3 Å and indicates the intramolecular hydrogen bond. This arrangement is supported by the conjugated π- electron system.

Generally, the tautomeric equilibrium of the isolated molecules might be different in solution since the solute-solvent interaction may stabilise the tautomers differently. The dipole moments in table 2 give a first hint on the preferred occurrence of the oxo-amino tautomer in polar solvents; the solubility of the hydroxo-imido form might be promoted in less polar solvents.

UV/VIS spectroscopy of ACV and DCV dissolved in methylene chloride give almost identical spectra. According to Clark and Tinoco[52] the signals at 300 nm can be assigned to isolated $n \rightarrow \pi^*$ transitions whereas $\pi \rightarrow \pi^*$ transitions occur approximately at 245 nm exhibiting a shoulder at the red-shifted side.

Albinson and Norden[53] found that insertion of a methyl group does not significantly influence the absorption signals. Conclusively, we would assume that both compounds have similar π-electron systems in the apolar methylene chloride. Hence the OH function in ACV may not affect the energetic transition observed at this wavelength. Most likely what is seen in the spectra are the electronic transition of the hydroxo-imido tautomer of ACV.

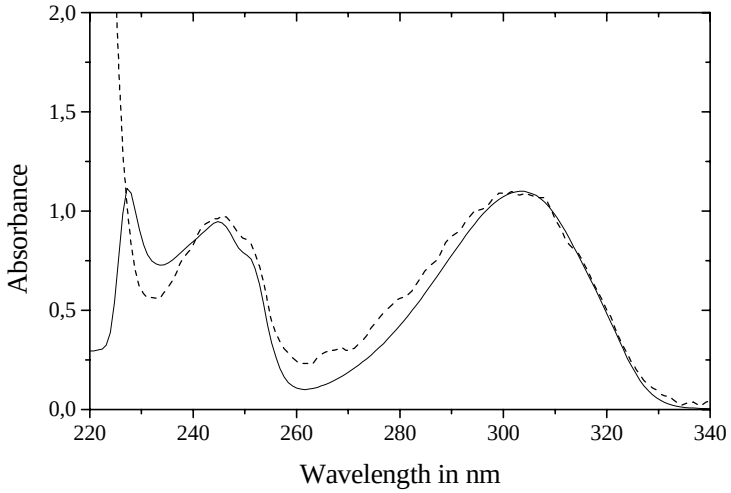


Fig. 3: UV spectra of ACV (dashed) and DCV (solid) dissolved in methylene chloride

Different observations were received for ACV if more polar solvents like acetonitrile and water are used. In acetonitrile a band maximum at 276 nm appears whereas the signal at 252 nm is almost unchanged in its position; the band at 301 nm vanishes. Contrary, the signals of DCV are almost unchanged. Hence, the shift of the electronic transition observed for ACV cannot be traced back to a solvatochromic effect. The spectra of the aqueous ACV solution combines the observations of the signals in methylene chloride and in acetonitrile. Based on the semi-empirical calculation and the assignment in acetonitrile and in methylene chloride the observed signals indicate the equilibrium of oxo-amino and hydroxo-imido tautomers in aqueous solutions whereas either the oxo-amino or the hydroxo-imido tautomer are found in acetonitrile and methylene chloride, respectively. The acyclovir derivatives investigated in this study did not show similar effects. In the case of the N-acetylated species the intramolecular hydrogen bond stabilises the oxo-amino form also in apolar solvents.

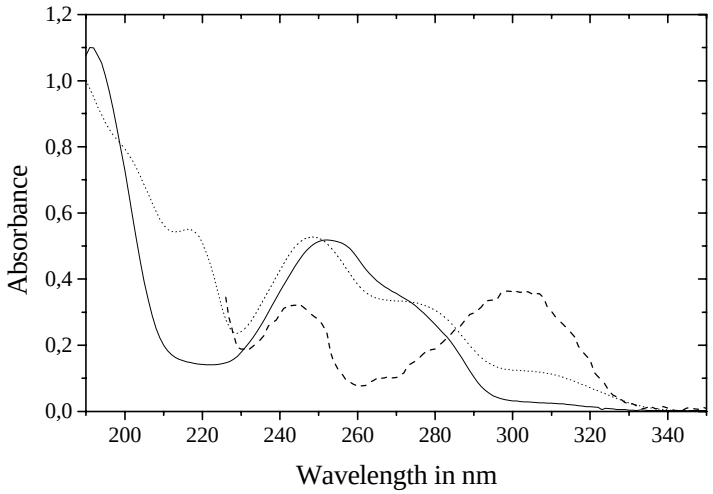


Fig. 4: UV spectra showing the solvent effect on ACV; (dashed line: CH_2Cl_2, dotted line: H_2O, solid line: AcN)

^{1}H NMR investigation allows an approximation of the concentration ratio of both tautomers. In agreement with the recent paper by Gao and Mitra[54] the assignment of the signals can be done unambiguously. Again ACV behaves differently than its derivatives. Thus, the C-8-H singlet and the N-9-CH_2O singlet occur twice in dependence on the concentration. In saturated D_2O the concentration ratio is 1:1. Obviously, at higher concentration the aggregate formation forces the shift of the tautomeric equilibrium towards the hydroxo-imido form.

Table 3: ^{1}H NMR signals of the DCV and ACV derivatives in solution in D_2O [in ppm]

Assignment	DCV	ACV	NAcACV	diAcACV
C**H$_3$**CO (O-ac)				1.793 (s)
C**H$_3$**CO (N-ac)			2.183 (s)	2.186 (s)
C**H$_2$**OH	3.569 (s)	3.569 (s)	3.577 (s)	
C**H$_2$**O				3.772 (m)
C**H$_2$**OC(O)				4.064 (m)
NC**H$_2$**O		5.403 (s)		
	5.542 (s)	5.506 (s)	5.516 (s)	5.514 (s)
C-8-**H**		7.817 (s)		
	8.192 (s)	8.119 (s)	8.035 (s)	8.039 (s)
C-6-**H**	8.625 (s)			

s=singlet, m=multiplet

These spectroscopic and semi-empirical calculations explain why the observed and calculated water-solvent partition coefficient for acyclovir differ so strongly. Thus, ACV itself can exist in different tautomeric forms in different solvents. Hence the determination of logP values will be difficult because the partitioning coefficient of the reference molecule ACV will refer to some mixture of both tautomers in aqueous media whereas the acetyl derivatives refer to the oxo-imido tautomer only.

The acyclovir derivatives substantiated that structural peculiarities might cause difficulties in the prediction of logP data. It is, therefore, the task of this work to determine the descriptors of oligopeptide derivatives based on the solvation equation by Abraham (eq. 4) and to discuss those with the results of spectroscopic, mainly FTIR, data on the same compounds. The effects of peptide enlargement, conformational flexibility of the peptide backbone, the influence of associative interactions as well as the behaviour of the diastereoisomers will be to the fore of our investigation.

B RESULTS

B.1 QSPR studies on oligopeptides based on HPLC measurements

B.1.1 Methods of analysis

It was the intention of these investigations to deduce solute descriptors for peptide compounds recognising also the diastereoisomeric properties. Up to our knowledge it is the first attempt to receive descriptors based on the solvation equation by Abraham eq. 4 for this class of compounds. The determination followed the creation of an equation system of five equations with three unknown parameters, the descriptors π_2^H, $\sum\alpha_2^H$ and $\sum\beta_2^H$. R_2 and V_x can be calculated individually based on atom and bond properties. From the theory it would have been sufficient to develop a equation system based on three solute properties. However, to reduce effects caused by specific interactions we set up the following equation.

$$\begin{pmatrix} \log SP_1 \\ \log SP_2 \\ \log SP_3 \\ \log SP_4 \\ \log SP_5 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{pmatrix} + \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{pmatrix} R_2 + \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{pmatrix} \pi_2^H + \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} \sum \alpha_2^H + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{pmatrix} \sum \beta_2^H + \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \\ \nu_5 \end{pmatrix} V_x \qquad [5]$$

In preliminary investigations it is necessary to characterise the partitioning systems by their coefficients c, r, s, a, b and ν. Furthermore, the choice of the partitioning systems determines the precision of the descriptors and hence requires special attention. With logSP data of the peptides and the coefficients the solution of eq. 5 can be received by a fitting approach using the sum of the squares of the differences between calculated and experimental logSP data to find the best fitting values for π_2^H, $\sum\alpha_2^H$ and $\sum\beta_2^H$.

Now the question arises which solute properties SP to take as parameters in the equation. It was our interest to develop a method which can be used as a general tool for the determination of descriptors. Traditionally the logarithm of the equilibrium concentration of a compound in two immiscible liquids, the partitioning coefficient P, represents a solute related property logSP which can be received from the shake-flask method. Among these logP values the logarithm of the equilibrium concentration in octanol-water, $\log P_{Oct}$ is most widely used. However, this procedure is quite time-consuming and demands a certain amount of the substance. Hence it would be helpful if this parameter can be approximated.

Collander had first demonstrated that there exists a linear relationship between the logarithmic values of partitioning coefficients of different partitioning systems[55]. Also the limitations of the relationship were shown which could be surmounted by the extensive work by Leo who stated more precisely that this correlation is applicable for compounds of related structure or partitioning systems of similar properties[56].

The determination of partitioning coefficients does not meet the requirements of high throughput techniques, however, the retention factor of a compound (logk) in reversed phase high performance liquid chromatography (RP-HPLC) is related to the logarithm of the distribution coefficients of a substance between mobile and stationary phase. Hence it can be taken as a measure for the hydrophobicity as well and has the advantage of a smaller material consume and the automation of the process [57, 58, 59, 60].

The idea arises to develop an approach for the determination of descriptors based on HPLC measurements which are performed under gradient conditions. Hence some preliminary experiments were necessary. Firstly, it needed to be proved that data received from gradient experiments are able to characterise the hydrophobicity of a compound and thus correlate with logP data. Secondly, the best gradient conditions need to be found with high chromatographic selectivity and high throughput. Furthermore, five appropriate chromatographic systems need to be selected which exhibit coefficients of high variety. The investigations are performed using compounds with known descriptors from the UCL database. However, peptides may behave differently in gradient experiments because they are able to form a high variety of conformers and can change their interacting surface area which leads to differences in the retention behaviour. Therefore, it should be tested whether the gradient conditions can be also applied on peptides before developing descriptors from these measurements. The details of the preliminary experiments are discussed in the next paragraphs.

B.1.1.1 Gradient experiments for the determination of CHI

It was mentioned above that the retention factor of a compound logk in reversed phase HPLC can be used as an alternative to logP. However, certain precautions must be taken when using $logk_W$[61] because the solute factors which influence $logP_{Oct}$ are not the same as those which influences logk' or $logk_0$[62, 63] Therefore, Valko and Slegel[64] introduced a new chromatographic hydrophobicity index φ_0 which can be calculated from $logk_0$ and includes the slope S of the regression line in the isocratic experiment according to eq. 6. The physical meaning of the chromatographic parameters S and $logk_0$ can be understood as follows: The S

values are related to the hydrophobic contact area between solute and the n-alkyl ligands whereas $\log k_0$ values are a measure of the affinity of the interaction under aqueous conditions[65, 66].

$$\varphi_0 = -\frac{\log k_0}{S} \qquad [6]$$

The index φ_0 represents the volume of the organic phase required to achieve a 1:1 distribution between mobile and stationary phase and ranges between 0 and 100. In terms of the isocratically determined $\log k_0$ this would correspond to the conditions were logk is equal to zero. The different φ_0 values for the chromatographic systems indicate the specific physical or chemical interactions of the alkyl aryl ketones with the mobile phase, bonded phase and support material. Table 4 exemplifies the φ_0 values of a series of homologues extended by paracetamol and acetanilide in order to extend the φ_0 scale to the low range region. According to Smith[67, 68] these homologues fulfil the requirements for HPLC retention indices[69] and are, therefore, defined as the standard set for calibration purposes.

Table 4: Isocratically determined φ_0 values for the standard mixture of five different columns

	$\varphi_{0\,CN}$	$\varphi_{0\,IAM}$	$\varphi_{0\,CD}$	$\varphi_{0\,Pro}$	$\varphi_{0\,In}$
Paracetamol	-1.8	2.9	11.8	6.3	19.9
Acetanilide	4.7	11.5	32.8	42.4	42.8
Acetophenone	15.4	17.2	46.9	64.0	71.5
Propiophenone	25.5	25.9	53.4	74.4	80.5
Butyrophenone	32.8	32.0	56.7	81.2	84.4
Valerophenone	38.5	37.3	60.4	86.7	91.2
Hexanophenone	42.2	41.8	64.2	91.2	95.3
Heptanophenone	44.1	45.6	67.5	95.7	98.8
Octanophenone	46.1	49.4	70.5	99.7	101.8

The comparison of the data in the table shows the effect of the stationary phase on the retention of a compound with constant mobile phase (acetonitrile/ammonium acetate buffer, pH 7.4). For the determination of φ_0 value for a compound the performance of measurements of several solvent compositions is necessary which is in contrast to the requirements of high throughput technologies. Recently, Valko et al.[70] reported about a new chromatographic hydrophobicity index (CHI value) based on fast gradient HPLC measurements which can be

used as an alternative to $logk_0$ and $logP_{Oct}$. Under certain conditions (flow rate, pore size, column length, gradient time etc.) the retention time of standard compounds of the gradient experiment are linearly correlated to the φ_0 value. Using the experimental conditions shown below the correlation between the φ_0 values given in table 4 for the octadecyl silicate phase (Inertsil column) is depicted in Fig. 5.

Experiment	
0.0-0.5 min	0% organic modifier
0.5-4.0 min	100% organic modifier
4.0-5.0 min	100 % organic modifier
5.0-5.2 min	0% organic modifier
5.2-7.5 min	0% organic modifier

The time intervals with constant solvent composition were inserted in order to allow a calibration of the system.

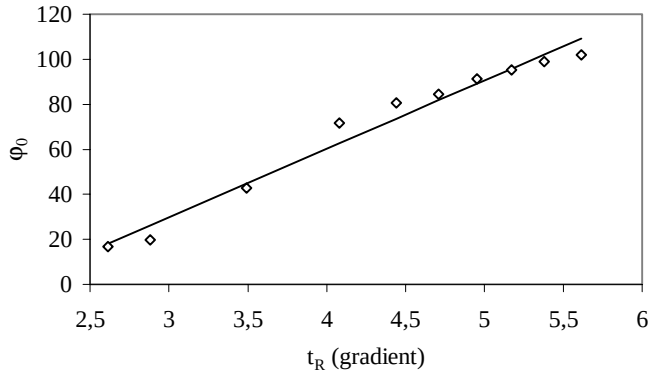


Fig. 5: Correlation of φ_0 vs. t_R of the standard mixture measured with the fast gradient experiment (0 to 100% in 3.5min) on the example of the ODS2 -IK (Inertsil) column

As can be seen in the plot the major problem arises from the change of the solvent strength which results in the curving. Kuronen[71] has explained that a linear relationship of φ_0 vs. t_R is very difficult to obtain for a series of homologous compounds in gradient elution without using complicated, non-linear gradient profiles. From the tested systems, the ODS2 columns show the largest deviation from linearity in comparison to the other columns. This seems to be a peculiarity of the octadecyl bonded phases and is based on the pocket formation into which the molecule diffuses[72]. Consequently, the adjustment of the equilibrium between the mobile phase and the stationary phase might take longer than for the other phases. Experiments with longer gradient times were not satisfying because peak tailing was observed for the

compounds with longer elution times. For the sake of equal treatment all experiments were performed using the same time table for all studies.

In Fig. 5 also the linear regression line is plotted. The linear regression between the φ_0 values of the standard set (S) and the retention time t_{RS} of the gradient experiment gives the parameters A and B (eq. 7). These parameters can be now used to calculate the chromatographic hydrophobicity index (CHI) for any compound X from their retention times t_{RX} measured under gradient conditions (eq. 8).

$$\varphi_0 = A \cdot t_{RS} + B \qquad [7]$$

$$CHI_X = A \cdot t_{RX} + B \qquad [8]$$

The linear regression parameters used for the calcualtion of the CHI values are given in paragraph E.2, table A2.

CHI can be treated analogously to φ_0 expressing also the percentage of organic modifier in the mobile phase for achieving a distribution equals to 1 between mobile and stationary phase. The CHI value opens the way to high throughput screening since it can be received from linear gradient measurements within a few minutes depending on the column properties to respond the change of the organic phase composition[73].

B.1.1.2 Selection of chromatographic systems for the descriptor determination

It was shown how the experimental conditions for the determination of CHI are chosen. Now the next step is the search for chromatographic distribution systems. The selection is directed on finding systems with most different coefficients c, r, s, a, b and ν. Hence 20 reversed-phases HPLC systems were tested with 30 structurally unrelated compounds. The descriptors of the 30 test compounds[74] and the details of the chromatographic systems are given in tables E1-E3.

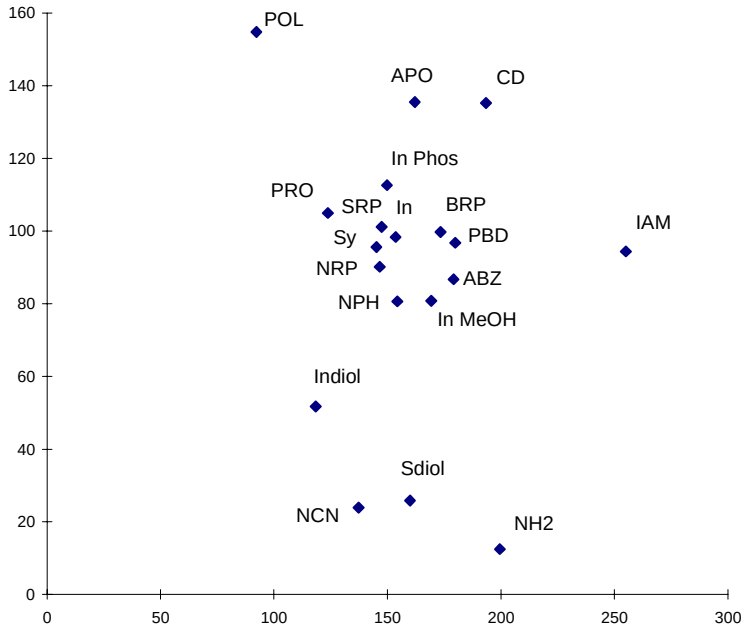


Fig. 6: The non-linear map of column principle component loadings (Abbreviations: mobile phase: acetonitrile/ NH_4OAc buffer, Columns with C18 alkyl chains: **APO**, **PRO**, **SRP**, **BRP**, **In**, **NRP**, **ABZ**, **SY**, **PBD**, with phenyl residues: **NPH**; with polar groups: **Indiol**, **Sdiol**, **NH2**, **NCN**, **IAM**; Polymer matrix*:* **POL**; β-permethylated cyclodextrine matrix*:* **CD**;
Silicate columns with C18 alkyl chains, mobile phase: acetonitrile/phosphate buffer, **InPhos**, mobile phase: methanol/NH_4OAc buffer: **InMeOH**. (for further details see: table E2 and E3)

Fig. 6 depicts the column principle component loadings for the tested chromatographic systems. The graphical plot shows that the silicate phases with apolar functions on the surface show similar retention behaviour. Hence we have chosen the five chromatographic columns listed in table 5 which exhibit most different parameters. As a mobile phase acetonitrile/0.05 molar ammonium acetate buffer (pH 7.4) was used. Elution properties can be also influenced by changing the organic modifier. These effects are described later on the hand of the peptide derivatives.

Table 5: Coefficients of the chromatographic systems based on multiple regression analysis of the CHI values of 30 test compounds (CHI values in table A3)

System	r^2	r	s	a	b	v	c
IAM	0.972	10.2 ±2.7	-11.0 ±2.9	6.5 ±2.8	-47.4 ±3.5	44.0 ±2.4	0.7
CD	0.970	7.5 ±3.8	-4.2[a] ±4.2	-1.9[a] ±4.0	-52.0 ±5.0	31.5 ±3.4	36.5
CN	0.957	9.0 ±4.1	-13.1 ±4.5	-7.1 ±4.2	-30.0 ±5.4	48.8 ±3.7	-18.0
Pro	0.993	3.4 ±2.3	-12.4 ±2.5	-23.2 ±2.4	-61.9 ±3.05	58.1 ±2.1	39.8
In	0.987	5.9 ±1.8	-15.3 ±2.0	-19.2 ±1.9	-63.7 ±2.4	65.0 ±1.6	28.6

[a]not significant

Thus, the procedure for the descriptor determination for peptide derivatives can be done. The descriptors R_2, V_x π_2^H, $\sum\alpha_2^H$ and $\sum\beta_2^H$ are the unknown parameters. But only the dipolarity/polarisability π_2^H, the effective hydrogen bonding acidity $\sum\alpha_2^H$ and the effective hydrogen bond basicity $\sum\beta_2^H$ are terms determining interactions between solute and solvent and, thus, can be received only from the experiment. The molar excess refractivity R_2 and the characteristic McGowan volume V_x can be calculated from the refractive index[75] and from the molecular structure based on the atom volumes and the bond constants[76], respectively. Therefore, the CHI values for the peptides on each chromatographic system are determined based on the parameters of the linear regression of the retention time of the standard set and their φ_0 values. Those CHI values were then used as solute related property in the set of linear equations as it was demonstrated by eq. 5. The solving the equation system in [9] by minimisation of the sum square of the observed and calculated CHI values gives finally the descriptors for the peptide derivatives.

$$\begin{pmatrix} CHI_{IAM} \\ CHI_{CD} \\ CHI_{CN} \\ CHI_{In} \\ CHI_{Pro} \end{pmatrix} = \begin{pmatrix} c_{IAM} \\ c_{CD} \\ c_{CN} \\ c_{In} \\ c_{Pro} \end{pmatrix} + \begin{pmatrix} r_{IAM} \\ r_{CD} \\ r_{CN} \\ r_{In} \\ r_{Pro} \end{pmatrix} R_2 + \begin{pmatrix} s_{IAM} \\ s_{CD} \\ s_{CN} \\ s_{In} \\ s_{Pro} \end{pmatrix} \pi_2 + \begin{pmatrix} a_{IAM} \\ a_{CD} \\ a_{CN} \\ a_{In} \\ a_{Pro} \end{pmatrix} \sum\alpha_2^H + \begin{pmatrix} b_{IAM} \\ b_{CD} \\ b_{CN} \\ b_{In} \\ b_{Pro} \end{pmatrix} \sum\beta_2^H + \begin{pmatrix} v_{IAM} \\ v_{CD} \\ v_{CN} \\ v_{In} \\ v_{Pro} \end{pmatrix} V_x \quad [9]$$

In the introduction we have discussed the coefficients for biological systems. For the determination of the CHI values we have used a immobilised artificial membrane carrying

lecithin functions on the surface of the stationary phase matrix. It seems worth discussing the coefficients of the chromatographic systems in comparison to those of the biological systems.

According to the solvophobic theory[77, 78, 79] the isocratic retention factor k' is related to the overall difference between mobile and stationary phase of the unitary GIBBS free energy ΔG° of the separation process. Hence eq. 10 correlates the retention factor k' and the GIBBS energy:

$$\ln k' = \ln\left(\frac{RT}{PV}\right) - \frac{\Delta G^{\circ}}{RT} + \Phi \qquad [10]$$

with Φ representing the functional relationship to the system phase ratio. The terms V and P are herein the main molar volume of the solvent and the operating pressure, respectively.

The overall change of the unitary GIBBS energy is strongly dependent on the specific interaction of the solute during the chromatographic separation. Indeed, ΔG° can be expressed as a composite of differences in free energies between mobile and stationary phase, associated with formation of cavity of molecular dimensions of the solute (ΔG°_{cav}), electrostatic interactions (ΔG°_{es}), van der Waals interactions (ΔG°_{vdw}), selfassociation or heterogeneous association in the absence of the solvents, e. g. in the gas phase ΔG°_{ass}. $\Delta G^{\circ}_{red.}$ covers the solvent-solute interactions which are not represented in the first three terms of eq. 11 and can be understood as a correction factor for non-linear behaviour.

$$\Delta G^{\circ} = \Delta G^{\circ}_{cav} + \Delta G^{\circ}_{es} + \Delta G^{\circ}_{vdw} + \Delta G^{\circ}_{ass} + \Delta G^{\circ}_{red} \qquad [11]$$

Depending on the chromatographic system the GIBBS energy terms play different roles. In the reversed-phase experiment where mainly partitioning effects are involved in the retention, the cavity formation term is essential.

Similar properties as in eq. 11 are also represented in the solvation equation (eq. 4). Here the coefficients r, s, a, b, ν and c describe the portions of the specific interaction on the retention. Hence the coefficient of the volume term, ν, is directly related to the capability of the cavity formation of the molecular dimensions of the solute by representing the volume influence. The comparison of ν of the ODS columns (In and Pro) carrying octadecyl chains on the surface with that of the IAM and CN which have phospholipids and propylcyanol functions bonded on the silicate matrix, respectively, shows that in the former system the

volume term is more important. Although the negative sign for s and b symbolises the affinity to reduce retention by stronger polar and hydrogen bond interactions with the mobile phase the magnitude of these coefficients show that in the IAM and CN column polar interactions between solute and stationary phase are also involved. It should be noted here as well that both ODS columns behave similarly which is represented in the coefficients. Small deviations are caused by the differences in the loading of the anchor molecules.

Table 6: Standardised coefficients of the chromatographic systems in comparison with other biological systems

System	r/v	s/v	a/v	b/v
IAM	0.231	-0.249	0.147	-1.077
CD	0.236	0.134	-0.060	-1.648
CN	0.184	-0.268	-0.145	-0.614
Pro	0.058	-0.212	-0.399	-1.064
In	0.091	-0.235	-0.303	-0.979
glycol/heptane	-0.036	-0.576	-1.716	-1.056
BBB	0.19	-0.69	-0.72	-1.28
skin/water	0.00	-0.327	-0.354	-1.949
SDS micelles/water	0.194	-0.143	-0.0476	-0.565

When these investigations were started it was hoped to find a chromatographic system with similar properties as the partitioning in biological systems. Especially the immobilised artificial membrane seemed to have the potential to simulate membrane properties since the lecithin functions on the silicate surface should result in comparable molecular interactions[80]. Pidgeon and Venkatarum demonstrated that the packing on the silicate surface is comparable with the molecular packing in membrane bilayers[81]. But also the immobilised artificial membrane column does not give similar properties as biologically relevant distribution systems nor do the other tested systems (Table 6). However, as Abraham et al.[38] have pointed out with known coefficients of the biological system and the descriptors of the prodrug determined forehand logBB can be approximated.

B.1.1.3 Peculiarities of peptide investigations using gradient experiments

The measurements described so far were done on model compounds from the UCL database which consists with some exceptions of small molecules with rigid structure. The title compounds are peptide derivatives of two to three amino acid units where the

conformation is not constrained by any tertiary structure. For larger peptide molecules Hearn et al. have given numerous examples like bombesin, β-endorphin[65], glucagon[82, 83], neuropeptides[84] and insulin[85] which exhibit a change in the interaction of the peptide with the hydrophobic surface area of RP phases in dependence on the experimental conditions. Hence the retention time and the bandwidth of the chromatographic signal were affected in a different manner as for molecules with hindered conformational flexibility by the choice of the gradient time. In order to apply the general method of CHI determination on oligopeptides with three amino acids which are not restricted to defined conformation but might have a certain distribution of the conformers, relevant parameters of the experimental set-up should be checked for peptide derivatives. These are

- the effect of the temperature on the retention time and bandwidth in isocratic conditions
- the effect of gradient time on the retention time e. g. CHI value,
- the effect of solvent polarity of the organic modifier

B.1.1.3.1 Effect of the temperature

The change of the interacting surface area with the stationary phase can be monitored on the hand of a sudden change in the retention behaviour of the solute. Hence if our peptide molecules show any conformational changes with the temperature it should reflect on the retention time e.g. the retention factor $\log k_0$, the slope S and in the bandwidth w of the chromatographic signal. Measurements at 25 and 40 °C under isocratic conditions were performed using a ODS column and acetonitrile/ammonium acetate buffer (pH 7.4).

Table 7: Temperature dependence on the retention factor $\log k_0$, slope value S and hydrophobicity index φ_0

	25 °C				40 °C			
	$\log k_0$	S	r^2	φ_0	$\log k_0$	S	r^2	φ_0
Ac-Ala-OMe	0.193	-0.0069	0.996	28.03	0.614	-0.0189	0.957	32.47
Ac-D-Ala-OMe	0.192	-0.0069	0.996	27.77	0.585	-0.0181	0.951	32.30
Z-Ala-Ala-OMe	1.788	-0.0296	0.991	60.39	1.769	-0.0299	0.989	59.17
Z-Ala-Ala-Ala-OMe	1.651	-0.0296	0.986	55.79	1.289	-0.0209	0.956	61.68
Ac-Phe-OMe	1.459	-0.0256	0.984	56.98	1.452	-0.0261	0.980	55.62
Z-Phe-OMe	1.229	-0.0217	0.980	56.62	1.267	-0.0227	0.975	61.68
Z-Phe-Phe-OMe	3.362	-0.0429	0.997	78.36	3.291	-0.0428	0.996	76.88
Z-Phe-Phe-Phe-OMe	3.752	-0.0467	0.993	80.34	3.075	-0.0387	1.000	79.47
Ac-Val-OMe	0.947	-0.02	0.947	47.33	1.020	-0.0223	0.986	45.75
Z-Val-Val-OMe	2.722	-0.0368	0.996	73.96	2.677	-0.037	0.997	72.34
Z-Val-Val-Val-OMe	2.751	-0.0374	0.995	73.55	2.784	-0.0384	0.994	72.49
Z-Val-D-Val-Val-OMe	2.803	-0.0372	0.995	75.36	2.848	-0.0385	0.994	73.96

Measurements were only performed at two different temperatures which do not allow any detailed interpretation of the molecular flexibility of the compounds itself but considering the effects with increasing chain length several conclusions can be drawn.

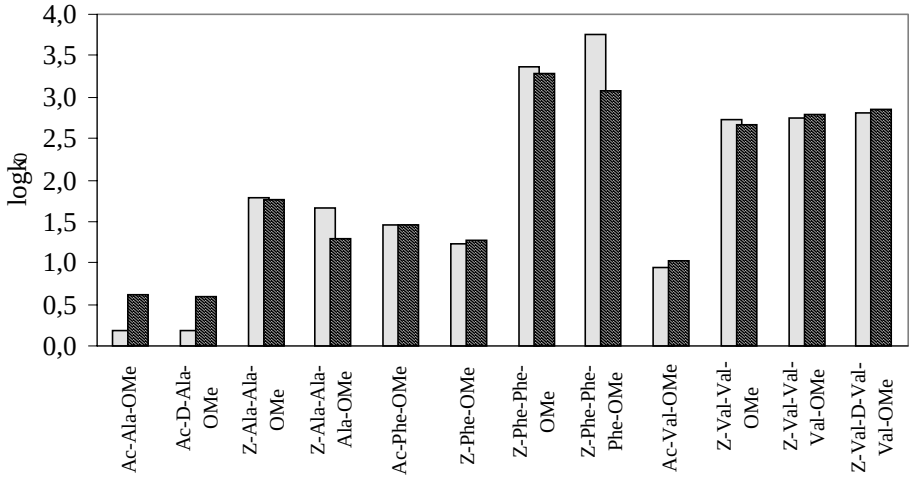


Fig. 7: Temperature effect on the retention factor $\log k_0$ for amino acid and peptide derivatives (grey: 25 °C, black: 40 °C)

In general, the effects in this temperature range are rather small. However, considering them more detailed it is obvious that an increase of the Brownian motion does not necessarily increase the retention factor. Furthermore, there is no clear dependence on the chain length and the increase in $\log k_0$. In some cases it rather seems that the contact time with the

stationary phase increases at higher temperatures as shown for Z-Phe-Phe-Phe-OMe. Hence the bulky residues might be more flexible and, thus, enable a more selective interaction with the stationary phase. It should be mentioned here that the small amino acid enantiomers of Ac-Ala-OMe do hardly interact with the reversed phase and, hence, the determination of $\log k_0$ is based on a small range of organic modifier concentration.

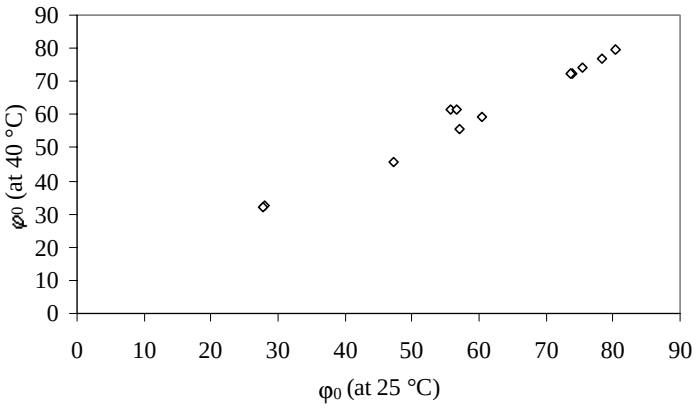


Fig. 8: The influence of the temperature effect on φ_0 for amino acid and peptide derivatives measured isocratically on an ODS column using acetonitrile/ammonium acetate buffer (pH7.4)

A different representation of the temperature effect is chosen in Fig. 8. Here it should be depicted that the increase of molecular flexibility does not affect the chromatographic hydrophobicity index φ_0 dramatically.

There exists a linear correlation between the φ_0 values measured at 25 and 40 °C. Deviations from this correlation are found for Ac-Ala-OMe, Ac-D-Ala-OMe, Z-Phe-OMe and Z-Ala-Ala-Ala-OMe which have a larger change in φ_0 values at 40 °C than the other substances. The retention times of the amino acid derivatives are very short at low concentration of the organic modifier and, therefore, hardly to determine precisely.

B.1.1.3.2 Effect of the gradient time

The temperature effect was studied under isocratic conditions but for the sake of higher efficiency the gradient experiment was chosen for the descriptor determination. For simple molecules with a rigid molecule moiety it was proved that the gradient time is restricted only by the parameters of the column to adapt on the change of the solvent polarity and its ability to change its solute properties. The influence of gradient time was tested on the same ODS column system using a flow rate of 1 ml/min. It was shown by Mant and Hodges[86] that the retention of synthetic peptides increases linearly with the reciprocal of the gradient slope

under ideal conditions which can be used as a working tool in the prediction of peptide retention.

Table 8: CHI values calculated from retention times at 25 °C of gradient HPLC measurements at different gradient times t_G [in min]

		CHI at different gradient time t_G								
Substance	φ_0	3.5	5.5	7.5	10	15	30	45	60	90
Ac-Ala-OMe	28.0	-	-	28.8	28.8	30.4	34.4	36.6	38.1	37.6
Ac-D-Ala-OMe	27.8	-	-	28.8	28.8	30.5	34.4	36.6	38.2	37.5
Z-Ala-Ala-OMe	60.4	53.8	55.3	55.5	56.3	57.8	60.2	61.5	62.2	60.1
Z-Ala-Ala-Ala-OMe	55.8	49.2	51.0	51.4	52.5	54.5	57.7	59.3	60.4	58.4
Ac-Phe-OMe	57.0	50.1	51.1	51.1	51.6	53.0	55.4	56.6	57.4	55.9
Z-Phe-OMe	56.6	48.6	49.6	49.2	49.3	50.5	52.7	53.7	55.2	53.3
Z-Phe-Phe-OMe	78.4	75.3	77.4	78.9	80.3	81.7	83.8	84.9	86.0	84.4
Z-Phe-Phe-Phe-OMe	80.3	78.0	80.2	82.0	83.7	85.4	87.7	89.2	90.2	88.7
Ac-Val-OMe	47.3	42.3	42.4	41.5	41.6	43.0	45.7	47.1	48.3	46.7
Z-Val-Val-OMe	74.0	69.5	71.2	72.1	72.9	74.1	75.8	76.9	77.8	75.8
Z-Val-Val-Val-OMe	73.5	69.0	71.0	72.1	73.2	74.7	76.7	78.0	79.0	77.2
Z-Val-D-Val-Val-OMe	75.4	71.2	73.0	74.0	75.0	76.3	78.2	79.5	80.4	78.6

Based on the linear correlation of the retention time of the standard alkyl aryl ketones and their φ_0 values the CHI values of the peptide derivatives could be derived. If the peptide derivatives are able to correspond simultaneously on the variation of the solvent polarity the calculated CHI values should be identical with the φ_0 values.

As it is shown on the example of the tripeptide derivatives the CHI values do not vary significantly with the gradient time. They are calculated in the range of the standard deviation of the isocratically determined φ_0 which was added at the zero value in the plot. As can be seen in all cases CHI calculated from the retention time of the 3.5 and the 5.5 gradient experiment is lower than φ_0. Two experimental reasons can be named as an explanation for this observation: Firstly, the retention time cannot be determined as accurately with such a fast gradient. Although the experiments were repeated at least two times the retention time varies especially in the cases of molecules (Z-Ala-Ala-Ala-OMe) with expected lower interaction with the stationary phase. Hence our experimental set-up is a compromise between high throughput method and accuracy of the CHI determination.

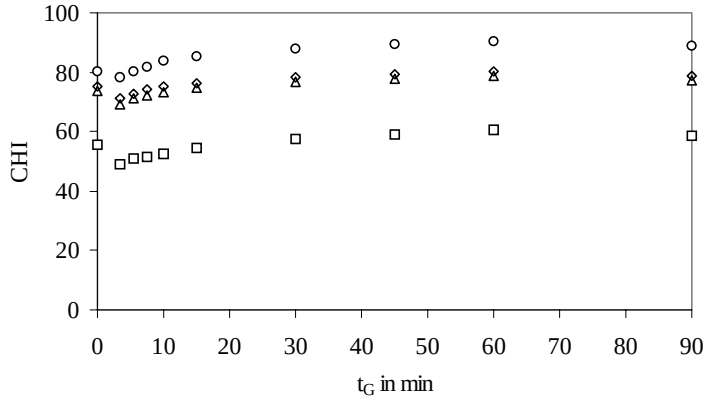


Fig. 9: The influence of the gradient time on the CHI value of peptide derivatives (square: Z-Ala-Ala-Ala-OMe, circle: Z-Phe-Phe-Phe-OMe, triangle: Z-Val-Val-Val-OMe, diamond: Z-Val-D-Val-Val-OMe)

As the result of the linear relationship between the gradient steepness and the retention time and consequently of the uniformity of the CHI values in Fig. 9 the theoretical treatment of Snyder[87], the linear solvent strength (LSS) gradient model, applies here. Hence the same chromatographic variables relevant to retention, resolution and bandwidth in isocratic elution also govern the gradient experiment.

B.1.1.3.3 Effect of the bandwidth

For studies of the conformation distribution the bandwidth of the chromatographic signal is a sensitive parameter. If UV array detectors are used in HPLC generally several absorption maxima were chosen for the observation. Hence the signals at 210, 230 and 245 nm were taken as an intensity measure. However, the absorption maximum at 210 nm is still strongly influenced by the absorption signal of the solvent. Ac-Ala-OMe does not absorb at 245 nm.

Therefore, further interpretations are based on the UV absorption signal at 230 nm which can be assigned to an electronic transition of the amide function sensitive to changes of the conformational flexibility of the peptide backbone. For dipeptide and tripeptide derivatives the 245 nm absorption gives identical results with regards to the distribution of conformers.

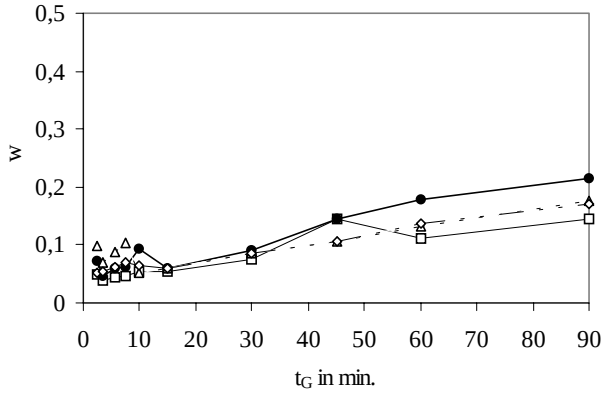


Fig. 10: Bandwidth w at 230 nm in dependence of the gradient time (square: Ac-Phe-OMe, circle: Z-Phe-OMe, triangle: Z-Phe-Phe-OMe, diamond: Z-Phe-Phe-Phe-OMe)

Analogue experiments on penta-L-phenylalanine have shown that when the gradient duration was increased larger increases at the bandwidth were observed at constant temperature[85]. However, in contrast to peptides with complex tertiary structure no significant increase was found which should indicate the change of the interacting surface area. The substances plotted in Fig. 10 behave similarly to the unprotected pentapeptide: There occurs a small change of the bandwidth with the gradient time which indicates the conformational dynamics at the hydrophobic surface but their are rather insignificant. Conclusively for the choice of the experimental conditions for the determination of chromatographic hydrophobicity indices the change in the conformation especially at fast gradients should not be of great importance. Note that the bandwidth in the case of Z-Phe-OMe indicates the highest variety of conformers.

B.1.1.3.4 Effect of solvent polarity

Generally, peptides at most pH values are charged molecules. The presence of counter ions and pH influences the retention. Hence for ODS columns perfluorinated carboxylic acids, particularly trifluoroacetic acid, TFA, and heptafluorobutyric acid, HFBA, are recommended by several authors as excellent components in organic modifiers since they are completely volatile, and allow UV detection at wavelength below 220 nm[88, 89, 90, 91, 92]. However, most of our chromatographic columns cannot stand these low pH values. Furthermore, these strong hydrogen bond donors are known to change the tertiary structure of the proteins[93] and the conformational equilibrium in smaller peptides as will be proved later. It was our interest to calculate structural descriptors based on the gradient retention times calibrated on the standard set of alkyl aryl ketones. As it was shown above the variation of the gradient time does not

have a significant influence on the retention time and also the effect on the bandwidth e.g. conformer distribution can be neglected. In paragraph B.1.1 it was explained how descriptors can be received from HPLC measurements. In order to select particular interactions as a relevant property in the retention process the column material was changed whereas the mobile phase was always acetonitrile/ammonium acetate buffer. As a second approach it would be possible to vary the organic modifier and leave the column system constant. Hence it can be assumed that purely partitioning would be the major separation phenomena. The amino acids and peptide derivatives are homologues whereas the side function and the number of hydrogen bond acceptor and donor functions characterises the molecules. By variation of the mobile phase the properties of the chromatographic column change. Furthermore, the solvation of the peptide molecules depends on the properties of the solvent as well. It is expected that the retention time will be a function of the mobile phase but based on the similarities of the peptide compounds the hydrophobicity should change linearly. Fig. 11 depicts the retention time under the conditions of the variation of the organic modifier. As a measure of the hydrophobicity the φ_0 values of the acetonitrile/ammonium acetate buffer are plotted on the y- axis. The gradient experiment was performed on shorter column systems (column length: 5 cm) and a 2.5 min gradient to enhance the throughput.

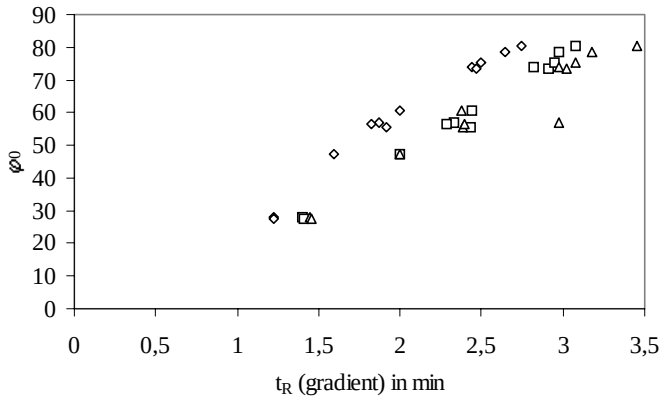


Fig. 11: Effect of organic modifier variation on the retention time of oligopeptide derivatives on an ODS column (organic modifier: square: MeOH, triangle: TFE, diamond: AcN)

The almost linear correlation between t_R of the aqueous methanol phase and φ_0 indicates that the sort of interaction of the solute and the stationary phase did not change drastically. In comparison to acetonitrile as organic modifier the curve is shifted. For TFE the linearity does not turn out satisfactorily. The deviation is drastical, especially in the case of Ac-Phe-OMe. Obviously, by the change of the solvation strength of the mobile phase the retention time becomes longer. It might be caused by the changed conformation as a consequence of

hydrogen bonding with the acidic trifluoroethanol. The CHI values will represent this unexpected behaviour of Ac-Phe-OMe as well, however, when five CHI values are combined for the fitting of the descriptors conformational aspects cannot be recognised and are averaged. When the stationary phase carries polar functions similar effects might become relevant.

Our investigations have shown that it is generally not advisable to use different organic modifiers for the descriptor determination of peptide derivatives because the different solvent interactions with the solute can cause strong standard deviations.

B.2 CHI values and descriptors of peptide derivatives

B.2.1 Effect of peptide enlargement on CHI and the descriptors

B.2.1.1 Effect of peptide enlargement on the hydrophobicity represented by CHI

According to the approach described above the CHI values of amino acids and oligopeptide derivatives with the hydrophobic amino acid residues alanine, phenylalanine, glycine and valine were calculated and summarised in table 9.

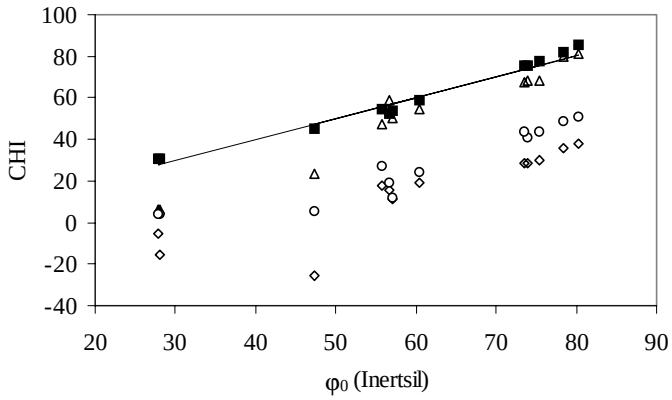


Fig. 12: Correlation of the CHI values of different chromatographic systems with isocratic chromatographic hydrophobicity index φ_0 (Inertsil) (squares: CHI_{In}, circles: CHI_{CN} triangles: CHI_{CD}, diamonds: CHI_{IAM})

In Fig. 12 the isocratic and the gradient chromatographic hydrophobicity index for peptide molecules measured on the chromatographic systems are depicted. The linear regression gives a residual standard deviation of 0.97 for the ODS column systems. As indicated by the solid line, the slope and the intercept are not exactly equal to 1 and zero, respectively, but this was not unexpected considering the influence of the gradient time discussed before. Generally, there are no changes in the type of interaction during the separation process for the homologue series of peptides. Stronger deviations are expected when the CHI values received from the

IAM, CN and CD column are plotted versus φ_0 (Inertsil). Although the separation process for the first two systems is mainly determined by partitioning polar interactions become important as well. The separation mechanism of permethylated β-cyclodextrine phases[94] involves exclusion processes and is thus not purely hydrophobic.

Valko et al.[70] had pronounced that the CHI value represent the hydrophobic properties of solutes. However, the substances for which descriptors were available were simple organic compounds (Table E1). For peptides the conformational aspects depend on the chain length as well as of the amino acid composition. Under the premise of the prediction of retention behaviour on hydrophobic stationary phases based on the structure of the peptide researchers have undertaken several attempts to calculate hydrophobicity by simple summation of retention coefficients of amino acids. Guo and Mant[88, 95, 96] have claimed that this approach works sufficiently well for small peptides with ca. up to 15 residues but fails with larger peptides if no coefficients for the peptide chain length correction are introduced. It has to be emphasised that these approaches are based on assumption that the peptides adopt random coiled structure and thus lack of a unique conformation. It is well-known that proteins and polypeptides form a variety of secondary structures such as helices and β-sheets. Peptides with up to four amino acid units are not able to gain formation energy by forming those hydrogen bonds. However, stabilisation may be possible by the formation of turn structures which are well-known to occur in cyclic peptides and linear oligopeptides. In particular, these structures were observed in the absence of solvents with strong hydrogen bond capabilities as hydrocarbons and chlorinated methanes as will be shown in the spectroscopic part of this work. Similar apolar properties may be found in the octadecyl chain of the ODS column.

More detailed analysis of the chromatographic hydrophobicity indices with respect to the peptide chain enlargement can be done with the values in table 9. Additionally in the last column the φ_0 data determined form isocratic investigations using the ODS2 Inertsil column and AcN/buffer are given.

Table 9: CHI values for amino acid and peptide derivatives measured on different chromatographic columns[a]

Substances	CHI_{IAM}	CHI_{CD}	CHI_{CN}	CHI_{In}	CHI_{InMeOH}	CHI_{Pro}	φ_0
Ac-Ala-OMe	-15.3	6.2	4.0	30.7	32.7	53.1	28.0
Ac-D-Ala-OMe	-5.2	5.9	4.0	30.8	33.0	53.1	27.8
Z-Ala-Ala-OMe	19.1	54.3	24.0	58.6	67.7	71.1	60.4
Z-Ala-Ala-Ala-OMe	17.6	47.6	27.3	54.5	67.1	68.8	55.8
Ac-Phe-OMe	11.0	50.1	11.7	54.0	64.6	68.5	57.0
Z-Phe-OMe	15.3	58.9	19.2	52.7	64.1	68.4	56.6
Z-Phe-Phe-OMe	35.7	80.0	48.6	81.8	84.0	85.4	78.4
Z-Phe-Phe-Phe-OMe	37.7	81.0	51.1	85.6	87.4	87.4	80.3
Ac-Val-OMe	-25.2	23.7	5.5	45.1	54.7	62.4	47.3
Z-Val-Val-OMe	28.4	68.6	40.6	75.5	79.7	81.0	74.0
Z-Val-Val-Val-OMe	28.4	67.5	43.4	75.6	81.8	81.1	73.5
Z-Val-D-Val-Val-OMe	30.1	67.9	44.	77.6	83.4	82.2	75.4
Z-Gly-Gly-OMe	11.8	50.8	13.6	52.5	61.2	67.3	-

[a]If not indicated the acetonitrile/ammonium acetate buffer was used as the mobile phase. For CHI_{InMeOH} the organic modifier was methanol.

Hydrophobicity scales for amino acids deduced in different possible ways use the amino acid with the smallest side chain, glycine, as the standard residue. Hence the influence of the side group can be approximated from CHI of the dipeptides. The effect of a methyl group can be quantitatively expressed as the difference of $(CHI_{Z\text{-}Ala\text{-}Ala\text{-}OMe}\text{-}CHI_{Z\text{-}Gly\text{-}Gly\text{-}OMe})/2$ and would give a value of 3.05 (CHI_{In}), 3.25 (CHI_{InMeOH}), 1.9 (CHI_{Pro}). The influence of an isopropyl group can be numbered with 11.5, 9.25 and 6.85, for CHI_{In}, CHI_{InMeOH} and CHI_{Pro}, respectively. In an analogous way the benzyl residue of phenylalanyl is approximated as 14.65, 11.4 and 9.05, for CHI_{In}, CHI_{InMeOH} and CHI_{Pro}, respectively.

If the amino acid contribution is additive for oligopeptides the addition of the corresponding residue value should give CHI of the tripeptide derivatives reduced by a constant value for the enlargement of the peptide backbone. Performing the simple arithmetic procedure gives the data in table 10.

Table 10: Difference between experimental and calculated CHI values when enlarging the peptide backbone

	ΔCHI_{In}	ΔCHI_{InMeOH}	ΔCHI_{Pro}
Ala	-7.15	-3.85	-4.2
Val	-11.40	-7.15	-6.75
Phe	-10.85	-8.00	-7.05
Phe[a]	-25.30	-16.5	-15.00

[a] calculated from Z-Phe-OMe

Two conclusions become obvious: The magnitude of the side chain contributions depends on the experimental conditions but the tendency of the effects will be similar. Secondly, the enlargement of the oligopeptide cannot be treated as linearly based on using Gly as the reference fragment. Hence for the deduction of hydrophobicity indices of amino acids it is essential to define a set of peptides with identical chain length also for the tripeptide derivatives if the investigations are aimed on the development of descriptors. Furthermore, the amino acid glycine is known to have a special role in peptide sequences because it was often found as the „turning“ amino acid residue. Hence its conformational behaviour is different from the properties of other amino acids.

The use of a random sampling of peptides of varying chain length would result in wrong parameters compared with parameters received from peptide retention of fixed chain length. In the case of polypeptides the stabilisation of tertiary structures by strong intramolecular hydrogen bonds or sulphur bridges which occur if hydrophilic amino acids as cystein and glutamine are involved lead to errors with respect to retention behaviour derived from peptides which are unable to interact in similar ways.

B.2.1.2 Effect of peptide enlargement on the descriptors

As it was pointed out earlier the hydrogen bonding capability plays an important role for the transport of a peptide through biological membranes. Thus, the estimation of the hydrogen bonding parameters according to eq. 5 can be done. The molar excess refractivity R_2 is based on the refractive index which was calculated for the solids using the ACS Chemsketch software package (VS. 4.0). Also the McGowan volume V_x can be determined from the sum formula using bond distances and atom constants. Naturally, R_2 and V_x are not changed by the diastereoisomerism of the peptides. Hence the differences in the retention of the LLL and LDL diastereoisomer of Z-Val-Val-Val-OMe can be traced back to the hydrogen bond and the polar properties.

Table 11: Descriptors of amino acid, di- and tripeptide derivatives determined from CHI

Substances	M_M [g/mol]	Refractive Index	R_2	π_2	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x [10^{-2} l/mol]	S. D.
Ac-Ala-OMe	145.16	1.426	0.217	0.36	-1.007	1.280	1.144	4.31
Ac-D-Ala-OMe	145.16	1.426	0.217	0.03	-0.656	1.246	1.144	5.84
Z-Ala-Ala-OMe	308.33	1.518	0.991	4.49	-0.104	0.914	2.349	2.86
Z-Ala-Ala-Ala-OMe	379.41	1.521	1.141	5.40	-0.07	1.294	2.887	2.99
Ac-Phe-OMe	221.25	1.514	0.838	3.72	-0.41	0.66	1.752	1.79
Z-Phe-OMe	313.35	1.561	1.508	5.74	-0.24	0.85	2.418	3.43
Z-Phe-Phe-OMe	460.53	1.584	2.360	7.75	-0.06	1.16	3.564	5.15
Z-Phe-Phe-Phe-OMe	607.71	1.597	3.235	11.10	-0.20	1.63	4.710	2.73
Ac-Val-OMe	173.21	1.433	0.193	2.42	-1.55	1.08	1.426	3.36
Z-Val-Val-OMe	364.44	1.507	0.945	5.35	0.09	0.98	2.912	5.45
Z-Val-Val-Val-OMe	463.57	1.507	1.063	7.34	0.16	1.33	3.732	4.65
Z-Val-D-Val-Val-OMe	463.57	1.507	1.063	7.26	0.17	1.32	3.732	4.32
Z-Gly-Gly-OMe	280.278	1.527	1.027	4.54	-0.33	0.79	2.068	2.03

From the definition of the descriptors it does not seem very reasonable to find negative values for $\sum\alpha_2^H$ since they have no physical meaning. Also it is rather unexpected that the values of the descriptors $\sum\alpha_2^H$ and $\sum\beta_2^H$ do not increase proportionally with the number of hydrogen bond acceptor and donor functions. Several explanations might be given: Although we tried to find chromatographic systems which represent all different interactions in a sufficient manner the hydrogen bond acidity represented by the coefficient a of the chromatographic systems is not very selective. Generally, we would have expected $\sum\alpha_2^H$ values close to zero for all investigated compounds since amides are in comparison to their hydrogen bond basicity weak proton donor systems. Secondly, the range of $\sum\alpha_2^H$ of the compounds used for the determination of the coefficients describing the partitioning system is limited by their availability in the UCL database which covers a comprehensive number of different compounds but does unfortunately not contain any peptide molecules yet. Furthermore, especially in the case of the IAM column negative CHI values were found for the protected amino acid derivatives. Here the retention is almost not affected by the lecithin functions at the silicate surface and no selectivity was found. Conclusively, in those cases the CHI values are uncertain and the standard deviation cannot be narrowed. We need to point out that these difficulties are not due to the use of CHI values for the determination of the descriptors but were specific to the peptides. The work of Du et al.[97] has shown that CHI data can be used as solute properties and the calculated descriptors for compounds incorporated in

the UCL database give acceptable results. Furthermore, the attempt of Abraham et al.[98] to estimate the descriptors from partitioning data resulted in similar difficulties.

Hence, it seems legitimate to revise the $\sum\alpha_2^H$ and to recalculate the descriptors with the restriction on $\sum\alpha_2^H$ being equal or larger as zero.

Table 12: Revised descriptors for homopeptide derivatives [V_x in 10^{-2} l/mol]

Substances	R_2	π_2	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x	S. D.
Ac-Ala-OMe	0.217	0.00	0.00	1.10	1.144	9.20
Ac-D-Ala-OMe	0.217	0.00	0.00	1.10	1.144	10.72
Z-Ala-Ala-OMe	0.991	4.64	0.00	0.90	2.349	5.20
Z-Ala-Ala-Ala-OMe	1.141	5.57	0.00	1.28	2.887	5.21
Ac-Phe-OMe	0.838	3.62	0.00	0.62	1.752	4.86
Z-Phe-OMe	1.508	5.79	0.00	0.83	2.418	5.46
Z-Phe-Phe-OMe	2.360	7.71	0.00	1.16	3.564	5.20
Z-Phe-Phe-Phe-OMe	3.235	11.21	0.00	1.61	4.710	5.75
Ac-Val-OMe	0.193	1.37	0.00	0.91	1.426	7.57
Z-Val-Val-OMe	0.945	5.58	0.09	0.97	2.912	7.18
Z-Val-Val-Val-OMe	1.063	7.34	0.16	1.33	3.732	4.65
Z-Val-D-Val-Val-OMe	1.063	7.26	0.17	1.32	3.732	4.32
Z-Gly-Gly-OMe	1.027	4.35	0.00	0.77	2.0668	4.25

The database at UCL does not contain any data for peptide derivatives but the group contribution approach by Platts et al.[99] which uses modified fragment descriptors received from the descriptors in the database should be used for comparison with our data. This group contribution approach does not simple work on the addition of fragments but also recognises electronic effects due to the closest environment. The results of this analysis are shown in table 13.

Table 13: Calculated descriptors from the fragments [V_x in 10^{-2} l/mol][99]

Substances	R_2[a]	π_2	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x
Ac-Ala-OMe	0.45	1.33	0.22	1.04	1.1439
Ac-D-Ala-OMe	0.45	1.33	0.22	1.04	1.1439
Z-Ala-Ala-OMe	1.35	2.12	0.32	1.65	2.3486
Z-Ala-Ala-Ala-OMe	1.61	2.89	0.54	2.26	2.8868
Ac-Phe-OMe	1.07	1.76	0.22	1.09	1.752
Z-Phe-OMe	1.71	1.77	0.11	1.09	2.418
Z-Phe-Phe-OMe	2.59	2.97	0.32	1.75	3.564
Z-Phe-Phe-Phe-OMe	3.48	4.17	0.54	2.4	4.7102
Ac-Val-OMe	0.43	1.3	0.22	1.06	1.4257
Z-Val-Val-OMe	1.32	2.04	0.32	1.69	2.9122
Z-Val-Val-Val-OMe	1.56	2.77	0.54	2.31	3.7322
Z-Val-D-Val-Val-OMe	1.56	2.77	0.54	2.31	3.7322
Z-Gly-Gly-OMe	1.3	2.19	0.32	1.62	2.0668

[a]The excess molar refraction R_2 was calculated from fragmental contributions as well. Contrary the values in table 11 and 12 are calculated from the refractive index.

The comparison of the data of the successive tables show that the descriptors π_2^H and $\sum\beta_2^H$ from the CHI data are larger than the estimated ones from fragmental approach. Contrary $\sum\alpha_2^H$ gives values which are significantly larger than zero. At this point the discrepancies of the values in both tables are not satisfying. However, they might not be caused purely by the evaluation of the descriptors from our CHI values but also from the limitations of the group contribution approach which is lacking on appropriate fragment data for e. g. ureido functions. Although this method is a strong improvement compared to the Klopman[100] fragment it also neglects any electronic effects in the peptide backbone. Those effects might strengthen the overall hydrogen bond basicity and contrary compensate the hydrogen bond acidity properties of the molecules. Despite of the deviation in numbers the tendency of the descriptors is represented quite well.

B.2.2 Effect of amino acid substitution in tripeptide derivatives

B.2.2.1 Effect of amino acid substitution on CHI

The significance of hydrophobicity for the stabilisation of proteins and peptides demanded the reception of hydrophobicity data of the naturally occurring amino acids which were derived from the transfer between octanol/water and first reported by Tanford[101] and Ben-Naim[102] in 1980. Since then, over 200 sets of parameters have been reported as a measure of

the biochemical and biological properties among them more than 80 just describing hydrophobicity[103, 104]. The data received for the amino acids vary significantly depending on how they were received from partitioning measurements or from HPLC retention times. However, also the selection of the amino acids and peptides for the analysis determines the value of the hydrophobicity scales. It was found that there is a good correlation between the amino acid partitioning data between hydrophobic and aqueous media[105] and the HPLC data of the analogue systems[89]. However, Cys, Lys and Arg were already identified as behaving different if the partitioning experiment is performed with the acetyl amino acid amides[106]. Hence, the charged residues show strong effects depending on the counterions used in the experiments and on the pH values. For reliable results it is, therefore, essential to control the effects of pH and counterion, self association of free amino acids and the solubility in the organic solvent[106, 107]. Also the choice of the peptide used for the deviation of amino acid hydrophobicity scales from retention data and received from multivariable analysis of peptide and protein system lead to quite different results. Wilce et al.[108] pointed out the ranking of the 20 essential amino acids strongly depends on the number and kind of the proteins on which the multiple regression analysis is based as well as on the RP- columns and solvent composition used for the determination of the retention times. Furthermore, Wilce et al. [109] expressed the influence of the mobile phase and the column material on the outcome of the amino acid ranking. The paper of Meek et al.[110] showed that the analysis of the retention data of HPLC measurements on ODS2 columns for peptides leads to different results in dependence on the experimental conditions. Furthermore, the influence of the vicinity on the hydrophobicity data need to be considered when the sequence of the peptides is changed[111].

According to the described procedure the CHI values of tripeptide derivatives of the sort Z-Ala-Xaa-Val-OMe are determined. The results for five different chromatographic systems are summarised in table 14. In order to neglect ionisation of polar side functions also these groups are protected with benzyl residues.

Table 14: CHI values of selected tripeptide derivatives of the type Z-Ala-**Xaa**-Val-OMe (LLL diastereoisomers)

Peptide	CHI_{CD}	CHI_{IAM}	CHI_{CN}	CHI_{Pro}	CHI_{In}
Z-Ala-**Trp**-Val-OMe	44.9	28.9	44.4	64.2	61.1
Z-Ala-**Phe**-Val-OMe	41.8	25.9	43.9	67.1	63.8
Z-Ala-**Met**-Val-OMe	41.6	24.6	40.3	-	64.6
Z-Ala-**Leu**-Val-OMe	38.6	23.3	41.9	66.0	61.9
Z-Ala-**Tyr**-Val-OMe	37.5	23.7	39.0	54.5	50.6
Z-Ala-**Val**-Val-OMe	35.9	-	38.4	61.0	56.6
Z-Ala-**Ala**-Val-OMe	34.3	19.7	33.4	-	56.5
Z-Ala-**Gly**-Val-OMe	33.6	20.6	32.1	-	54.4
Z-Ala-**His**-Val-OMe	28.8	18.4	35.7	42.0	37.8
Z-Ala-**Glu**-Val-OMe	27.5	17.1	22.7	-	47.6
Z-Ala-**Ser**-Val-OMe	26.0	17.6	30.7	45.6	41.0
Z-Ala-**Asn**-Val-OMe	23.6	16.6	30.5	42.1	37.6
Z-Ala-**Tyr(OBzl)**-Val-OMe	50.3	32.2	48.8	77.1	76.2
Z-Ala-**Lys(NBzl)**-Val-OMe	44.2	26.6	45.4	67.0	64.4
Z-Ala-**Asp(OBzl)**-Val-OMe	44.2	27.7	45.5	70.6	68.5
Z-Ala-**Ser(OBzl**)-Val-OMe	42.9	27.8	44.8	70.1	67.8

By systematic variation of the amino acid Xaa in the tripeptide derivatives Z-Ala-Xaa-Val-OMe the changes in the CHI values can directly be traced back to the hydrophobicity of the particular amino acid. Hence, by regression analysis the contribution of the amino acid Xaa can be estimated and compared with 14 amino acid hydrophobicity scales in the literature. These amino acid hydrophobicity scales were chosen to encompass the range of methods that have previously been utilised to generate hydrophobicity values and also to be representative of the types of hydrophobicity scales which has been generally applied to aid the elucidation of proteins or functional features. The abbreviation of the authors in the brackets were used in table 15 and 16.

Whereas the scales of Nozaki and Tanford (*Noz*)[105], Zimmerman et al. (*Zimm*)[112], Pliska and Fauchere (*Fauch*)[106] and Bull and Breese (*Bull*)[113] were received from solubility or partitioning measurements of amino acids or amino acid derivatives. The scales by Clothia (*Cloth*)[114] and Janin (*Jan*)[31] were statistically received from measurements of the accessible surface area of each amino acid from the crystallographically determined tertiary structure of 12 proteins and the ratio of the buried amino acid to accessible molar fraction of each amino acid in 22 proteins, respectively. The authors of the scales 11 to 14 combined different experimental scales and estimated coefficients in order to describe the hydrophobicity of

certain proteins or peptides. Thus, Guy (*Guy*)[115] calculated a hydrophobicity scale from the range of the previously measured transfer energy for each amino acid based on partitioning experiments in water and a less polar phase. Von Heijne and Blomberg (*Heij*)[116] estimated the free energy of transfer of a single amino acid in a polypeptide from a random coil in an aqueous phase to a helical membrane-bound conformation. This approach allows the formation of hydrogen bonds and charge distribution. Wolfenden et al. (*Wolf*)[117] analysed the transfer behaviour of organic molecules which represented the side chains of the essential amino acids from water to the vapour phase. Also the scale of Eisenberg et al. (*Eisen*)[118] is a result of the combination of five different scales. Analogously to our measurements the last complex of hydrophobicity scales was received from HPLC data but experimental conditions and the selection of peptides and proteins changed with regards of number and chain length. Hence, Parker et al. (*Parker*)[89] investigated the elution time of synthetic octapeptides of the type Ac-Gly-Xaa-Xaa-$(Leu)_3$-$(Lys)_2$-amide with systematic variation of the amino acid Xaa using a ODS HPLC column and linear gradient conditions (mobile phase: $NaClO_4$/AcN/H_2O, pH 7.0). Also Meek (*Meek*)[119] used the regression analysis of RP-HPLC retention times of 25 peptides in order to calculate amino acid hydrophobicity coefficients which was extended a year later on 100 peptides by Meek and Rosetti (*MeekR*)[110] (ODS, mobile phase: $NaClO_4$, AcN/H_2O, pH 7.0) The most comprehensive scale contains the data of the regression analysis of retention times of 1738 peptides (up to 50 residues) and was published by Wilce et al. (*Wilce*)[120] (ODS, mobile phase: TFA/AcN/water, gradient-HPLC, pH 7.0).

The amino acid scales based on the CHI values measured on the five chromatographic systems were received from the regression analysis which assumes that the CHI values of a tripeptide Z-Ala-Xaa-Val-OMe is the sum of the contributions of the C- and N- terminal protecting functions and the three amino acid residues. Hence the regression analysis of the data in table 15 gives the values in table 16.

It would have been also possible to receive a hydrophobicity scale for the side chain residues by using the CHI values of Z-Ala-Gly-Val-OMe as a reference compound. The result would, however, only vary in value.

Table 15: Amino acid hydrophobicity contribution as calculated from CHI in comparison with literature scales

						experimental			statistical			combination			HPLC			
	CHI_{CD}	CHI_{IAM}	CHI_{CN}	CHI_{In}	Noz	Zimm	Fauch	Bull	Cloth	Jan	Guy	Heij	Wolf	Eisen	Parker	Meek	MeekR	Wilce
Ala	4.31	-4	3.42	13.69	1	0.83	0.31	-200	-0.29	0.3	0.1	2.9	1.94	0.25	2.1	0.5	1.1	0.06
Arg						0.83	-1.01	-120	-2.71	-1.4	1.9	-9.4	-19.9	-1.8	4.2	0.8	-0.4	-0.85
Asn	-6.33	-7.06	0.54	-5.23	-1.5	0.09	-0.64	80	-1.18	-0.5	0.48	-1	-9.68	-0.64	7	0.8	-4.2	0.25
Asp						0.64	-0.77	-200	-1.02	-0.6	0.78	-5.6	-10.9	-0.72	10	-8.2	-1.6	-0.20
Cys					0	1.48	1.54	-450	0	0.9	-1.42	-0.08	-1.24	0.04	1.4	-6.8	7.1	0.49
Gln					-1	0	-0.22	160	-1.53	-0.7	0.95	-0.52	-9.38	-0.69	6	-4.8	-2.9	0.31
Glu	-2.49	-6.59	-7.25	4.74		0.65	-0.64	-300	-0.9	-0.7	0.83	-0.4	-10.2	-0.62	7.8	-16.9	0.7	-0.10
Gly	3.67	-3.06	2.18	11.54	0	0.1	0	0	-0.34	0.3	0.33	1.9	2.39	0.16	5.7	0	-0.2	0.21
His	-1.14	-5.25	5.78	-5.12	1	1.1	0.13	-120	-0.94	-0.1	-0.5	-1.5	-10.3	-0.4	2.1	-3.5	-0.7	-2.24
Ile					5	2.52	1.8	2260	0.24	0.7	-1.13	4.4	2.15	0.73	-8	13.9	8.5	3.48
Leu	8.65	-0.36	11.97	19.05	3.5	3.07	1.7	2460	-0.12	0.5	-1.18	4.2	2.28	0.53	-9.2	8.8	11	3.50
Lys						1.6	-0.99	-350	-2.05	-1.8	1.4	-2.3	-9.52	-1.1	5.7	0.1	-1.9	-1.62
Met	11.7	0.94	10.35	21.75	2.5	1.4	1.23	1470	-0.24	0.4	-1.59	2.1	-1.48	0.26	-4.2	4.8	5.4	0.21
Phe	11.9	2.19	14	20.87	5	2.75	1.79	2330	0.24	0.7	-2.12	4.4	2.15	0.73	-9.2	13.2	13	4.80
Pro					1.5	2.7	0.72	-980	-0.9	-0.3	0.73	1.1		-0.07	2.1	6.1	4.4	0.71
Ser	-3.97	-6.06	0.72	-1.87	-0.5	0.14	-0.04	-300	-0.75	-0.1	0.52	0.36	-5.06	-0.26	6.5	1.2	-3.2	-0.62
Thr					0.5	0.54	0.26	-520	-0.71	-0.2	0.07	1.2	-4.88	-0.18	5.2	2.7	-1.7	0.65
Trp	14.96	5.18	14.49	18.22	6.5	0.31	2.25	2010	-0.59	0.3	-0.51	3.9	-5.88	0.37	-10	14.9	17	2.29
Tyr	7.56	-0.02	9.03	7.75	0.5	2.97	0.96	2240	-0.71	-0.2	-0.21	1.2	-4.88	-0.18	-1.9	6.1	7.4	1.89
Val	5.95	-0.3	8.42	13.76	3	1.79	1.22	1560	0.09	0.6	-1.27	3.9	1.99	0.54	-3.7	2.7	5.9	1.59

Table: 16: Correlation of hydrophobicity scales

				experimental			statistical			combination				HPLC			
	IAM	CN	In	Noz	Zimm	Fauch	Bull	Cloth	Jan	Guy	Hej	Wolf	Eisen	Parker	Meek	MeekR	Wilce
CD	0.940	0.781	0.8212	0.760	0.265	**0.866**	0.676	0.479	0.560	0.549	0.682	0.307	0.682	**0.812**	0.555	**0.828**	0.452
IAM		0.7988	0.671	0.810	0.156	**0.899**	0.748	0.275	0.498	0.528	0.637	0.124	0.543	**0.812**	0.662	**0.880**	0.507
CN			0.445	0.688	0.349	**0.891**	0.747	0.357	0.592	0.703	0.512	0.215	0.604	**0.888**	**0.838**	0.691	0.438
In				0.583	0.221	0.633	0.487	0.698	0.619	0.481	0.811	0.553	0.774	0.604	0.333	0.609	0.493

As shown in table 16 the amino acid hydrophobicity scales received from CHI values correlate well with the scale of Pliska and Fauchere[106] which based on the partitioning data of amino acids in octanol/water and again substantiates the fact that the CHI values can be used as an alternative for partitioning data also in the case of peptide derivatives. It is satisfying that a good correlation occurs for the scale of Parker et al.[89] who performed HPLC measurements on the synthetic octapeptide with systematic substitution of one amino acids in peptide with constant chain length and sequence. It is exactly this reason why the most comprehensive scale based on the retention data of 1738 peptides and proteins fails in the correlation with our scales. As it was stressed before chain length and sequence effects as well as hydrogen bonding and ion formation do influence the overall hydrophobicity represented by the retention time and lead to significantly different results in hydrophobicity coefficients of amino acids.

Table 17: Ranking of the CHI values in comparison to the literature hydrophobicity scales[a]

Xaa	Zimm	Fauch	Meek	MeekR	Parker	Wilce	Guy	$CHI_{IAM.}$	CHI_{CD}	CHI_{CN}	CHI_{In}
Trp	16	1	1	1	1	4	7	1	1	1	4
Phe	3	3	3	2	2	1	1	2	2	2	2
Met	9	6	7	8	5	12	2	3	3	4	1
Leu	1	4	4	3	3	2	5	6	4	3	3
Tyr	2	8	6	5	7	5	9	4	5	5	8
Val	6	7	8	7	6	6	4	5	6	6	5
Ala	11	10	13	10	9	14	11	8	7	8	6
Gly	18	13	15	12	15	12	12	7	8	9	7
His	10	12	16	14	9	20	8	9	9	7	11
Glu	13	16	20	11	19	15	17	11	10	12	9
Ser	17	14	8	19	17	17	14	10	11	10	10
Asn	19	16	11	20	18	11	13	12	12	11	12

[a] A ranking of 1 indicated the most hydrophobic amino acid while a ranking of 20 denotes the least hydrophobic amino acid. Because of the lacking of some values in the scales deduced from CHI the ranking ends with 12. In some cases identical values were found for two amino acids and were, therefore, given identical ranking numbers.

Table 17 shows the ranking of the amino acid hydrophobicity of different scales. In agreement with the correlation coefficients in table 17 the scales established by Pliska and Fauchere and Parker et al. give similar tendencies as found for our scales. Surprisingly the ODS column (In) shows strong deviation from other HPLC scales. Kaibara et al.[72] mentioned that octadecyl columns may show pocket formation which might be more disturbing for the determination of the retention time of small molecules than for proteins. Also Guo et al.[88]

reported about difficulties on these columns due to the occurrence of adsorption and desorption processes.

B.2.2.2 Effect of diastereoisomerism and amino acid sequence on CHI

Hydrophobicity effects of peptides involving D-amino acids are hardly discussed in the literature although it is a well-known fact that diastereoisomers and enantiomers can be separated on reversed-phase chromatographic column systems. Conclusively the properties of these compounds in the interaction with the column material must be different as well. Indeed the CHI values of LLL, LDL and DLL diastereoisomers are slightly different in dependence on the chromatographic system. Table 18 gives a selection of the investigated compounds.

Table 18: Effect of diastereoisomerism and sequence on the CHI values

Substance	CHI_{IAM}	CHI_{CD}	CHI_{CN}	CHI_{In}
Z-Ala-Phe-Val-OMe				
LLL	25.9	41.8	43.9	63.7
LDL	26.1	42.6	43.4	65.5
DLL	26.0	42.6	44.0	64.0
Z-Ala-Leu-Val-OMe				
LLL	23.3	38.6	41.9	61.9
LDL	24.4	38.9	42.2	64.0
DLL	23.7	38.8	41.9	62.5
Z-Ala-Val-Leu-OMe	23.5	38.8	41.6	62.1
Z-Ala-Val-Phe-OMe	26.1	41.6	42.6	62.2

Similarly the hydrophobicity e.g. the CHI value change when the sequence of the amino acids in tripeptide derivatives are changed. Hence the mass of the molecule and the volume which is considered as the major magnitude for the retention are constant and nevertheless the CHI values represent different affinity of the molecule to the column material.

B.2.2.3 Effect of amino acid substitution and diastereoisomerism on the descriptors

In the former paragraph we have found that small structural changes in the tripeptide derivatives can affect the hydrophobic character of the compounds. To illuminate the structural differences we have analysed the CHI values in terms of the solute descriptors. The result of the regression analysis of five CHI values for diastereoisomeric tripeptide derivatives are summarised in table 19. The amino acid residues are sorted according to their increasing

hydrophobicity which is represented by the CHI value determined at permethylated β-cyclodextrine phase.

The last row of table 19 give the standard deviation (S. D.) in errors of CHI units. In general, the error was less than 3 CHI units. Exceptions were only found in the cases of amino acid residues with polar side chains (e. g. Glu) and with large protecting groups (e. g. Arg(Tos)). It is assumed that ionisation or adsorption occurs in the elution process. The variation of the pH values might help overcoming the problems of ionisation in the case of Glu but unfortunately our chromatographic systems cannot stand high pH. The precision of the descriptors π_2^H, $\sum\alpha_2^H$ and $\sum\beta_2^H$ is strongly dependent on the range of the coefficients of the five chromatographic systems shown in table 5. In the former paragraphs we have described the details of the descriptor calculation and it was shown that the range of the coefficients **b** is large whereas the range for **a** was small. Hence, we can give the precision of $\sum\beta_2^H$ with ±0.01 units whereas the S. D. for $\pi_2^H$and $\sum\alpha_2^H$ is larger being ±0.02 and ±0.05, respectively.

Table 19: Descriptors of diastereoisomeric tripeptide derivatives Z-Ala-Xaa-Val-OMe [V_x in 10^{-2} l/mol]

Xaa	R_2	π_2^H	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x	S. D.
Tyr(OBzl)						
LLL	2.475	8.61	0.00	2.13	4.584	1.87
LDL		8.68	0.00	2.11		1.90
Trp						
LLL	2.663	7.05	0.00	2.05	4.006	3.57
LDL		7.28	0.08	2.03		0.53
Lys(NBzl)						
LLL	1.928	8.39	0.21	2.22	4.590	0.17
LDL		8.43	0.09	2.14		1.17
Asp(OBzl)						
LLL	1.868	7.10	0.06	2.04	4.133	0.30
LDL		7.13	0.05	2.03		0.28
Ser(OBzl)						
LLL	1.779	6.54	0.02	2.00	3.976	1.55
LDL		6.63	0.07	1.98		0.30
Arg(Tos)						
LLL	3.218	9.76	0.12	2.45	4.877	3.68
LDL		9.73	0.15	2.44		3.65
Phe						
LLL	1.807	6.13	0.10	1.96	3.776	0.77
LDL		6.16	0.12	1.94		0.43
DLL		6.17	0.09	1.94		0.82
Met						
LLL	1.490	5.86	0.16	1.84	3.614	2.42
LDL		5.59	0.20	1.93		0.77
Leu						
LLL	1.096	5.28	0.21	1.88	3.591	1.50
LDL		5.26	0.18	1.87		1.08
DLL		5.29	0.21	1.87		1.34
Tyr						
LLL	2.047	6.49	0.36	2.06	3.835	0.97
LDL		6.49	0.37	2.05		0.99
Val						
LLL	1.091	5.05	0.24	1.87	3.450	1.56
LDL		4.77	0.07	1.86		3.16
Ala						
LLL	1.123	4.59	0.22	1.75	3.169	1.51
Gly						
LLL	1.141	4.21	0.30	1.71	3.028	1.04
DLL		4.32	0.38	1.74		1.35
His						
LLL	1.884	5.54	0.56	2.15	3.596	2.02
Glu						
LLL	1.414	6.00	0.29	1.95	3.525	3.38
LDL		4.50	1.41	1.99		4.58
Ser						
LLL	1.339	4.51	0.47	1.96	3.227	1.10
LDL		4.47	0.50	1.95		1.15
Asn						
LLL	1.439	4.96	0.54	2.10	3.425	1.14
LDL		4.93	0.56	2.09		1.19

The graphic expression of the data in table 19 are shown in figures 13 and 14. Generally it was found that with decreasing hydrophobicity -the amino acid hydrophobicity decreases from the left to the right in the figures- the effective hydrogen bond basicity decreases as well. This tendency is also found for the term describing the compound`s capability for polar interaction, π_2^H (not shown).

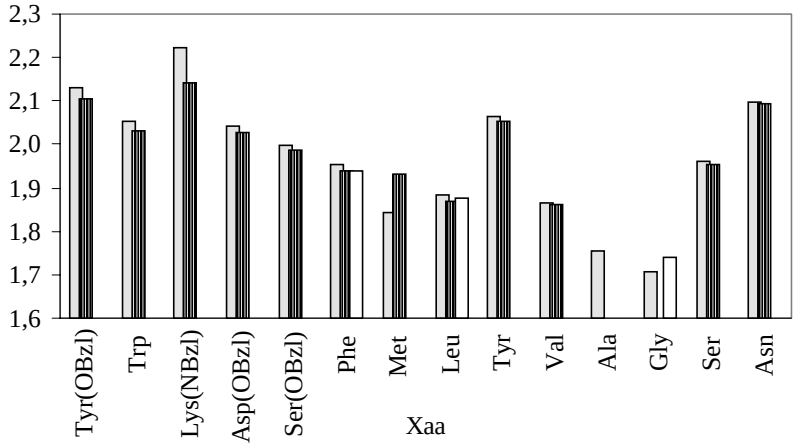


Fig. 13: Effective hydrogen bond basicity $\sum\beta_2^H$ of diastereoisomeric tripeptide derivatives Z-Ala-Xaa-Val-OMe (grey: LLL, black: LDL and white: DLL diastereoisomer)

Contrary the effective hydrogen bond acidity $\Sigma\alpha_2^H$ increases with decreasing hydrophobicity (Fig. 14). Furthermore, protection of the polar functions in the side chain results in an increase of the hydrogen bond basicity because by blocking of the OH and the NH functions they are converted into stronger hydrogen bond bases. Deviations from the relation between hydrophobicity and $\sum\beta_2^H$ (and π_2^H, respectively) were observed in the cases of Tyr, Ser and Asn as Xaa in the tripeptide derivative Z-Ala-Xaa-Val-OMe. This can be easily understood from the structure of these amino acids: Tyr and Ser carry OH groups in the side chain which can act as proton donors. Conclusively the basic properties of these derivatives might be compensated by the acidic character of the proton donor functions. The interplay of $\sum\beta_2^H$ and $\sum\alpha_2^H$ which are overall properties might result in the ranking of these derivatives in the hydrophobicity scale. Indeed the $\sum\alpha_2^H$ values for Tyr and Ser derivatives are surprisingly large.

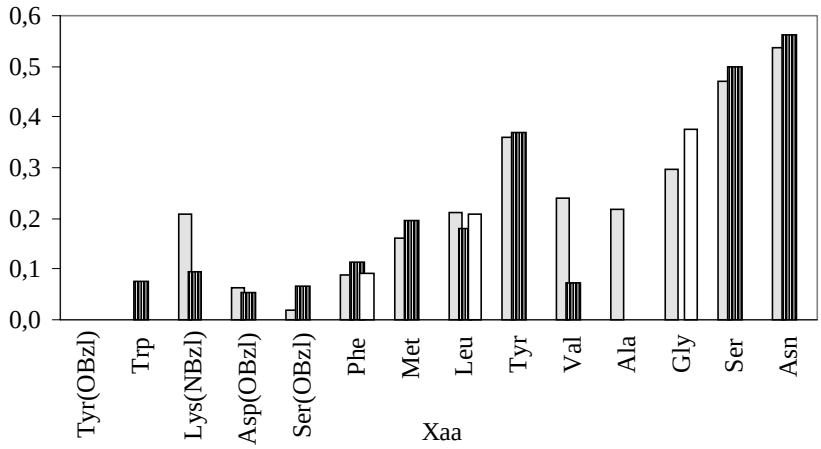


Fig. 14: Effective hydrogen bond acidity $\sum\alpha_2^H$ of diastereoisomeric tripeptide derivatives Z-Ala-Xaa-Val-OMe (grey: LLL, black: LDL and white: DLL diastereoisomer)

In the case of Tyr(OBzl) containing tripeptide derivative the high hydrogen bond basicity is not compensated by the hydrogen bond acidity. Conclusively, the hydrophobicity of this substance should be also large. Also Asn as Xaa in the tripeptide derivative contains an additional amide group in the side function and should, therefore, influence the overall basicity. Indeed the value for $\sum\beta_2^H$ is large for Z-Ala-Asn-Val-OMe which demonstrates impressively the analytical power of approach based on the solvation equation by Abraham.

Due to its definition the descriptors V_x and R_2 are not able to illuminate the differences in the CHI values of LLL, DLL and LDL diastereoisomers. Hence the differences in the hydrophobicity must be traced back to properties describing specific interaction of the chiral compounds with the stationary phase or to conformational effects in the peptides themselves. In the case of the hydrogen bond basicity it was clearly found that the LLL diastereoisomers exhibit slightly greater $\sum\beta_2^H$ values than the LDL analogue. Contrary the DLL diastereoisomers show similar values as LLL. We will take up the discussion of the conformational and association behaviour of the tripeptide derivatives in the spectroscopic part of this work. There it will be shown that the diastereoisomers behave differently in apolar environments which might explain the differences in $\sum\beta_2^H$. Unfortunately, the $\sum\alpha_2^H$ data do not allow such a differentiation.

B.2.3 Approximation of partitioning coefficients from solute descriptors

It was stated before that the knowledge of the descriptors of solutes have a great potential regarding approximation of any partitioning system for which the coefficients characterising the solvation system are known. It is generally accepted that a $\log P_{Oct}$ in the range of 2 is

optimal for a drug to pass to the active centres in the organism[121]. Hence in order to test if the tripeptide compounds would theoretically lie in the range of appropriate $logP_{Oct}$ values the coefficients of the solvation equation and the descriptors of the peptide derivatives were used to calculate the partitioning data in table 20.

Table 20: LogP values in octanol/water and blood/brain partitioning systems

Substances	$logP_{Oct}$	logBB
Ac-Val-OMe	1.04	-0.158
Z-Val-Val-OMe	2.50	-1.532
Z-Val-Val-Val-OMe	2.60	-2.205
Z-Val-D-Val-Val-OMe	2.72	-2.150
Z-Ala-Trp-Val-OMe	2.35	-1.801
Z-Ala-Phe-Val-OMe	2.27	-1.575
Z-Ala-Met-Val-OMe	2.18	-1.573
Z-Ala-Leu-Val-OMe	2.34	-1.338
Z-Ala-Tyr-Val-OMe	1.92	-1.971
Z-Ala-Val-Val-OMe	2.06	-1.341
Z-Ala-Ala-Val-OMe	1.92	-1.195
Z-Ala-Gly-Val-OMe	1.94	-1.099
Z-Ala-His-Val-OMe	1.61	-1.792
Z-Ala-Glu-Val-OMe	1.27	-1.943
Z-Ala-Ser-Val-OMe	1.63	-1.363
Z-Ala-Asn-Val-OMe	1.49	-1.603

From the logP values it can be deduced that tripeptide derivatives with polar side chains (Xaa=His, Glu, Ser, Asn) are not sufficiently hydrophobic whereas Z-Ala-Xaa-Val-OMe derivatives with bulky but apolar side chains as Trp, Leu and Phe might be dissolved in the hydrophobic regions of the membrane. Furthermore, the size of the molecule seems rather important since acetyl amino acid methyl esters might be washed out from the hydrophobic regions as well.

B.2.4 Fragmental descriptors

Besides the peptides discussed up to this point we have extended our analysis to tripeptide derivatives exhibiting a random amino acid sequence. However, in all cases the N- and C-terminal functions were protected with common functions known in peptide chemistry. Interactions of the polar head functions which would naturally occur and lastingly influence

the conformation can be neglected[122]. The central amino acid has as such an environment similar to this in a larger peptide or protein with no defined tertiary structure. We are aware of the fact that the application of our descriptors might lead to errors when they are used to predict the solubility and partitioning behaviour of molecules which do not fit into the homologues series of tripeptide derivatives with certain structural features. However, when QSPR methods are used the aim is not primarily focused on the estimation of accurate numbers of logP but to have a tool for the pre-selection of potential drugs. Consequently we have analysed the extended set of descriptors containing 61 tripeptide derivatives deduced from CHI values and used the multiple regression method to define values for amino acid residues. The regression bases on the assumption that the five descriptors are the sum over the contribution of each amino acid and the protecting groups at the C- and N- terminal end of the molecule. Hence the sequence and conformational effects are neglected. Peptides containing D-amino acids were treated differently from those consisting of L-amino acids. Besides the descriptors discussed in tables 12 and 19 the tripeptide descriptors in table A8 were used to receive the fragmental descriptors in tables 21, 22 and 23.

Table 21: Contributions of the amino acid residues to the descriptors of tripeptide derivatives [V_x in 10^{-2} l/mol]

Fragments	R_2	π_2^H	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x	Number of contributions
Gly	0.126	0.00	0.139	0.550	0.433	13
Ala	0.104	0.23	0.02	0.59	0.573	52
D-Ala	0.102	0.15	0.05	0.59	0.573	4
Val	0.076	0.88	0.00	0.62	0.855	62
D-Val	0.081	0.83	0.00	0.64	0.855	6
Leu	0.080	1.05	0.00	0.67	0.996	5
D-Leu	0.082	1.00	0.00	0.66	0.996	1
Phe	0.795	1.98	0.00	0.72	1.181	11
D-Phe	0.792	1.79	0.00	0.74	1.181	2
Pro	0.476	1.03	0.00	0.66	0.747	2

The regression was performed by a linear search of the minimum values of the square sum of the differences between calculated and experimental descriptors. Thus, by the definition of V_x and R_2 as a sum of the contribution of atomic and bond properties, the square sum was equal to zero. For the experimentally determined descriptors the deviation was larger and can be given with 0.75, 0.13 and 0.05 for π_2^H, $\sum\alpha_2^H$ and $\sum\beta_2^H$, respectively.

Table 22: Contributions of the amino acid residues with polar functions in the side chain to the descriptors of tripeptide derivatives [V_x in 10^{-2} l/mol]

Fragments	R_2	π_2^H	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x	Number of contributions
Ser	0.325	0.25	0.05	0.75	0.632	1
D-Ser	0.325	0.21	0.07	0.74	0.632	1
Tyr	1.033	2.06	0.09	0.89	1.240	2
D-Tyr	1.033	2.23	0.00	0.84	1.240	1
Met	0.476	1.59	0.00	0.63	1.019	1
D-Met	0.476	1.32	0.00	0.72	1.019	1
His	0.870	1.27	0.00	0.94	1.001	1
Trp	1.649	2.78	0.00	0.84	1.411	1
D-Trp	1.649	3.02	0.00	0.82	1.411	1
Asn	0.425	0.70	0.00	0.89	0.830	1
D-Asn	0.425	0.67	0.00	0.89	0.830	1
Glu	0.400	1.73	0.00	0.74	0.930	1
D-Glu	0.400	0.24	0.00	0.78	0.930	1
Ser(OBzl)	0.765	2.27	0.00	0.79	1.381	1
D-Ser(OBzl)	0.765	2.36	0.00	0.78	1.381	1
Tyr(OBzl)	1.461	4.35	0.00	0.92	1.989	1
D-Tyr(OBzl)	1.461	4.36	0.00	0.90	1.989	1
Lys(NBzl)	0.914	4.24	0.00	0.93	1.995	1
D-Lys(NBzl)	0.914	4.17	0.00	0.93	1.995	1
Asp(OBzl)	0.854	2.84	0.00	0.83	1.537	1
D-Asp(OBzl)	0.854	2.87	0.00	0.82	1.537	1
Arg(Tos)	2.204	5.47	0.00	1.24	2.282	1
D-Arg(Tos)	2.204	5.50	0.00	1.25	2.282	1

Due to the inhomogenous set of available tripeptide derivatives the contributions of the amino acids occurring only once in the structure must be treated with precaution. However, the values for a number of amino acids with apolar side chains as Ala, Phe, Val, Leu should be representative. Interestingly the polar interactions represented by π_2^H correlate with the size of the amino acid (Fig. 15). Contrary the overall hydrogen bond acidity $\sum\alpha_2^H$ and basicity $\sum\beta_2^H$ is almost independent of the type of the amino acid but determined by the number of acceptor and donor positions. This can be exemplified on the $\sum\beta_2^H$ values for amino acids with apolar side chains where $\sum\beta_2^H$ is calculated as 0.5 to 0.8 whereas Ser(OBzl), Tyr(OBzl), Lys(NBzl), Asp(OBzl) and Arg(Tos) exhibit higher values due to the additional presence of

electron lone pairs in the side chain. For the diastereoisomeric tripeptide derivatives an influence of D-amino acids on hydrogen bond basicity was observed. As expected this effect is not as significant anymore since the statistic neglects sequential and conformational peculiarities.

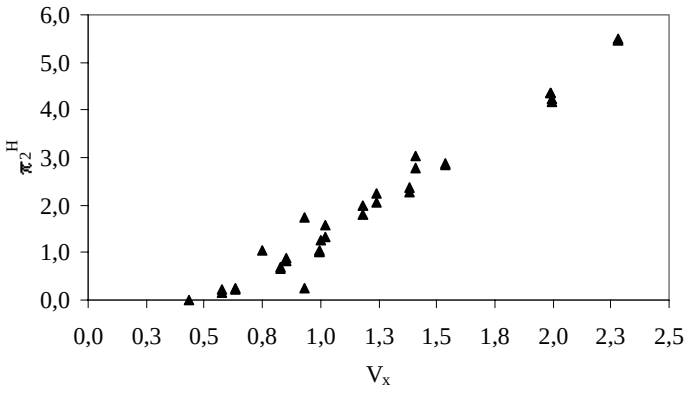


Fig. 15: Correlation between of V_x of an amino acid residue and its dipolar properties described by π_2^H

The protecting functions on the N- and C- terminal end need to be considered in the prediction of V_x, R_2 and π_2^H but can be ignored when hydrogen bond interactions are of interest since these values are zero.

Table 23: Fragmental descriptors of protecting functions [V_x in 10^{-2} l/mol]

Fragments	R_2	π_2^H	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x	Number of contributions
Boc	0.020	1.72	0.00	0.00	0.694	2
Z	0.725	2.64	0.00	0.00	0.879	59
OEt	0.102	0.86	0.00	0.00	0.428	4
OMe	0.109	0.51	0.00	0.00	0.287	53
OnBu	0.083	1.39	0.00	0.00	0.710	2
OBzl	0.783	2.29	0.00	0.10	0.895	2

B.2.5 Prediction of dipeptide descriptors from fragmental amino acid descriptors

In the former part it was always stressed that the accuracy of hydrophobicity parameters as well as descriptors is strongly a function of the chain length and the selection of peptides for the regression analysis. Now it should be shown to which extend the fragmental amino acid descriptors can be used for the prediction of properties of substances of similar structure but

consisting only of two amino acids and, therefore, exhibiting different conformational behaviour. Again we have deduced the descriptors for a number of dipeptides with mainly apolar side chains from the CHI values received by the well described high throughput HPLC method. In part these descriptors were already discussed in previous paragraphs. Additional CHI values are given in the appendix as well as the predicted descriptors. Due to the nature of V_x and R_2 the calculation of these descriptors from our fragment values gives the same values as calculated individually for each dipeptide.

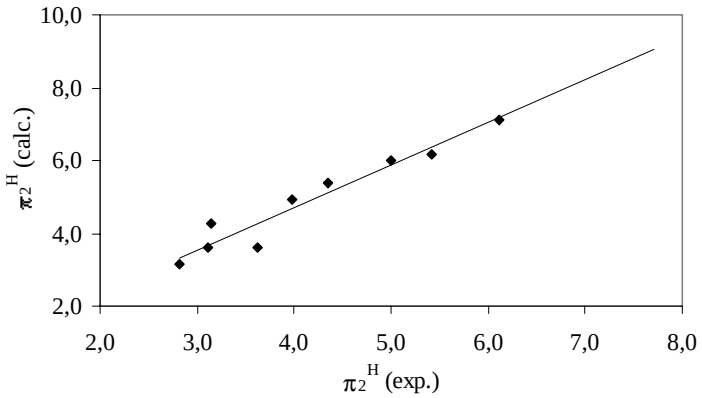


Fig. 16: Correlation of the predicted descriptor π_2^H with those determined from HPLC measurements (r^2=0.938)

As the figures 16 and 17 show there exists a good correlation between the predicted values and the descriptors determined according to our approach. But it should be emphasised that the slope of the correlation are only close to one and the intercept is not exactly zero. This seems to be a typical handicap of the fragmental data that within a homologous series of compounds a linear correlation is found but the precision is strongly determined from the selection of the compounds. The deviations in the plot of the $\sum\beta_2^H$ are caused by Z-Ala-Ser-OMe and Z-Ala-Trp-OMe.

As it would have been expected the predicted $\sum\alpha_2^H$ are not satisfying. However, up to our knowledge this is the first attempt to create such descriptors for amino acids and it will need an extended data set to improve the fragment descriptors in order to receive a potential method with predictive power for this class of compounds.

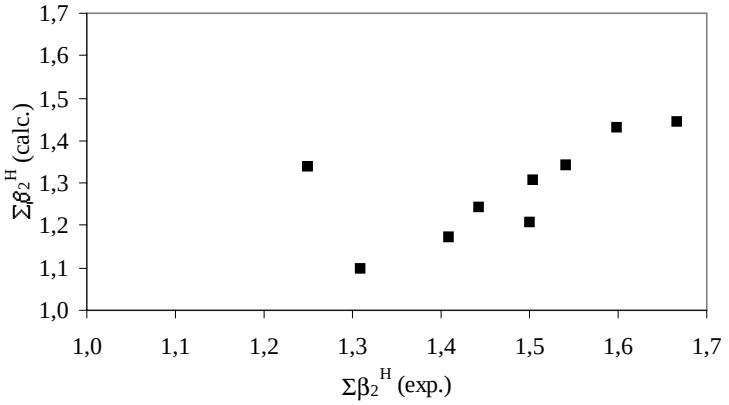


Fig. 17: Correlation of the predicted descriptors $\sum\beta_2^H$ with those determined from HPLC measurements (r^2= 0.9113)

In paragraph B.2.1.1. we have shown that the prediction of CHI for tripeptides by addition of a CHI fragment to a dipeptide CHI value does not give satisfying results. From these studies we have deduced that the size of the peptide molecules is an important parameter in the descriptor deduction. Furthermore, the choice of Gly as a reference molecule needed to be questioned.

B.2.6 Descriptors for ureido sugar amino-acid and dipeptide derivatives

We have shown that the fragmental descriptors can be used successfully for the prediction of dipeptide descriptors within a tolerance of 10 %. Still the structural pattern of the tripeptide and dipeptide derivatives are quite similar. Hence it was of interest to investigate a different class of peptide derivatives and to test the reliability of the fragmental descriptors.

Studies on N-(2-amino-2-deoxy-β-D-glucopyranoside)-N'-carbamoyl-L-amino-acid- and -dipeptidyl- esters were started because of their analogy to streptozotocin (2-deoxy-2-(3-methyl-3-nitrosoureido)-D-glucopyranose) [123], and chlorozotocin (2-deoxy-2-[(2-chloroethyl)-3-nitrosoureido]-D-glucopyranose)[124, 125]. Those drugs are proved to be effective anticancer drugs, show, however, strong side effects[126, 127]. In order to make the therapeutics more tolerated by the human body with remaining high effectiveness further research is focused on the exchange of the residues in the ureido side chain by linking body-own or body-like substances like peptides or amino acids to the streptozotocin moiety. As intermediates in the synthesis 3-nitrosoureido glucopyranoside derivatives the title compounds are formed.

Fig. 18 sketches the general ureido glucopyranoside structure where the amino acid and dipeptide residues are bonded. According to the three letter code recommended by IUPAC the amino acid and dipeptide residues are listed in table 24.

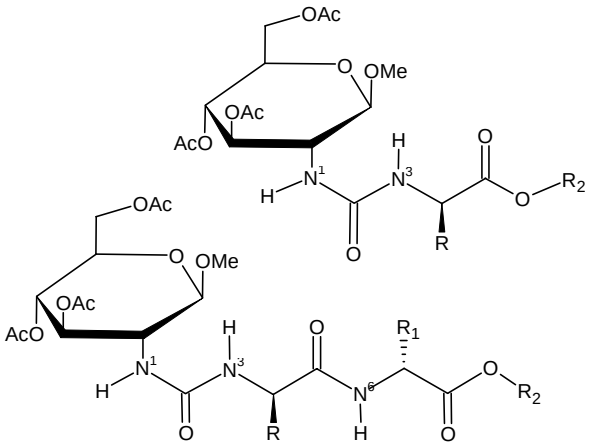


Fig. 18: Sketch of the acetylated N-(2-amino-2-deoxy-β-D-glucopyranoside)-N'-carbamoyl amino acid (above) and dipeptidyl derivatives; the residues R are listed in table 24.

Table 24: Residues of the amino acid and dipeptide derivatives linked to the ureido glucopyranoside moiety sketched in Fig. 18.

Abbreviation	Residue R	R_1	R_2
-Gly-OEt	H		CH_2CH_3
-Ala-OEt	CH_3		CH_2CH_3
-Val-OEt	$CH(CH_3)_2$		CH_2CH_3
-D-Val-OEt	$CH(CH_3)_2$		CH_2CH_3
-Leu-OEt	$CH_2CH(CH_3)_2$		CH_2CH_3
-Ile-OEt	$CH(CH_3)(CH_2CH_3)$		CH_2CH_3
-Phe-OMe	$CH_2C_6H_5$		CH_3
-Gly-Ala-OEt	H	CH_3	CH_2CH_3
-Ala-Gly-OEt	CH_3	H	CH_2CH_3
-Ala-Ala-OEt	CH_3	CH_3	CH_2CH_3
-Gly-Val-OEt	H	$CH(CH_3)_2$	CH_2CH_3
-Leu-Gly-OEt	$CH_2CH(CH_3)_2$	H	CH_2CH_3
-Gly-Phe-OBzl	H	$CH_2C_6H_5$	$CH_2C_6H_5$
-Phe-Gly-OEt	$CH_2C_6H_5$	H	CH_2CH_3
-Ala-Phe-OEt	CH_3	$CH_2C_6H_5$	$CH_2C_6H_5$
-Gly-Val-OBzl	H	$CH(CH_3)_2$	$CH_2C_6H_5$

According to the well-described procedure the descriptors of these compounds were deduced from the CHI values (Table A10) and compared with those calculated from the fragmental descriptors (Table A11). A selection of the descriptors is given in table 25.

On the hand of the graphical plots (Fig. 19 and 20) the results will be discussed concisely. As expected the McGowan characteristic volume V_x and the excess molar refraction R_2 can be deduced from the fragmental descriptors with high accuracy (r^2 value equal to 0.999) if the corresponding values for the acetylated N-(2-amino-2-deoxy-β-D-glucopyranoside)-N'-carbamoyl moiety are set equal to 2.289 and 0.706 for V_x and R_2, respectively. As stated before the effective hydrogen bond acidity $\sum\alpha_2^H$ cannot be obtained from the fragments. From the experimental CHI values $\sum\alpha_2^H$ was estimated as zero for all compounds with the exception of -D-Val-OEt. However, the amino acid and dipeptide derivatives contain two and three NH functions, respectively, which are classical proton donor functions. This fact should be represented by $\sum\alpha_2^H$ if not bulky residues makes the NH function inaccessible and lower the overall hydrogen bond acidity.

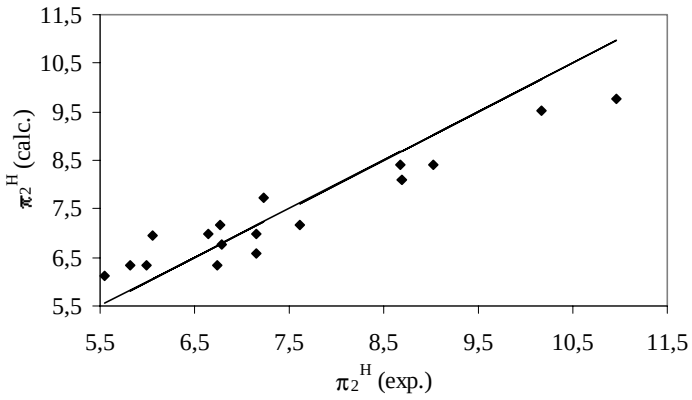


Fig. 19: Correlation of the experimental π_2^H values with those calculated from fragmental descriptors (π_2^H for the acetylated N-(2-amino-2-deoxy-β-D-glucopyranoside)-N'-carbamoyl residue: 5.25)

The term describing the solute's capability of polar interactions π_2^H correlate satisfactorily with the calculated values. The deviation from the predicted values is less than 10 %. It seems to be a general tendency to estimate smaller values when adding fragmental descriptors as it was discussed on the hand of table 12 and 13. Hence we tend to assume that this is a sign for the existence of preferred conformations.

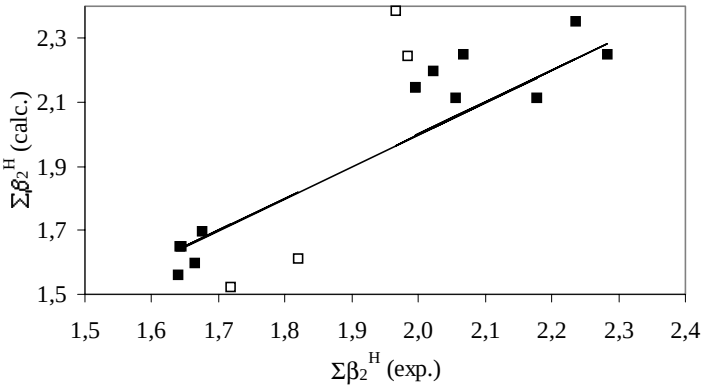


Fig. 20: Correlation of the experimental $\sum\beta_2^H$ values with those calculated from fragmental descriptors ($\sum\beta_2^H$ for the acetylated N-(2-amino-2-deoxy-β-D-glucopyranoside)-N'-carbamoyl residue: 0.975)

The effective hydrogen bond basicity $\sum\beta_2^H$ is determined by the number of effective acceptor functions in the molecule. Hence in the plot two areas can be spotted which are due to the amino acid and the dipeptide derivatives. The deviation between calculated and experimental values is within the limits of tolerance of 10 % for most substances. Among all derivatives the strongest differences between the values were observed for ureido sugars with -Gly-OEt, -D-Val-OEt, -Phe-Gly-OEt and -Ala-Phe-OBzl side chains with -11.2, -11.4, 13.3 and 21.5% deviation, respectively, and are marked in the graphical depiction. Again the spectroscopic investigation will help to understand these peculiarities and hence we refer to paragraph B.3.3

For the discussion of the sequence effect some data for compounds with different sequence are selected in table 25.

Table 25: Sequence effect on the descriptors of ureido derivatives

	experimental descriptors					predicted descriptors				
Compound	R_2	π_2^H	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x	R_2	π_2^H	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x
-Gly-Ala-OEt	1.114	5.82	0.00	2.18	3.705	1.038	6.34	0.24	2.11	3.723
-Ala-Gly-OEt	1.114	6.73	0.00	2.06	3.705	1.038	6.34	0.24	2.11	3.723
-Gly-Val-OBzl	1.886	9.03	0.00	2.07	4.453	1.691	8.42	0.15	2.25	4.472
-Val-Gly-OBzl	1.886	8.68	0.00	2.28	4.453	1.691	8.42	0.15	2.25	4.472

The effect of the sequence can be discussed on the descriptors π_2^H and $\sum\beta_2^H$. As expected the sequence influences the magnitude of the experimentally determined descriptors whereas the calculated descriptors cannot recognise sequence changes. For dipeptide residues this

effect is rather small; the values differ by less than ±8%. The descriptor π_2^H is mainly affected by the glucopyranoside and its polar acetyl functions. The $\sum\beta_2^H$ values for the acetylated N-(2-amino-2-deoxy-β-D-glucopyranoside)-N'-carbamoyl residue was estimated to 0.975 which is in the same range of magnitude as the contributions of the amino acid residues. However, from the data in table 25 no conclusions on the tendency of the deviation of -Gly-Xaa- or -Xaa-Gly- sequences can be drawn.

B.3 Spectroscopic studies on oligopeptide derivatives

B.3.1 Solute proton donor and proton acceptor scales

In the former chapter it was shown that the experimentally received descriptors π_2^H, $\sum\alpha_2^H$ and $\sum\beta_2^H$ play a key role in the quantitative structure-property relationships in model partitioning systems such as chromatographic systems. The descriptors can be used to predict the partitioning coefficient of a substance in any partitioning system and also in biological membranes either directly when the coefficients of the partitioning systems are known or via the correlation between logP, $\log k_0$ and CHI. However, the task for the medical chemist would be, once the descriptors of a prodrug are known, to enhance the solute related property by systematic molecular modification that may raise the activity relative to the previous expectations. This problem is not trivial because the enhancement of the hydrogen bond acidity by introducing an additional proton donor function might change the conformation simultaneously because of a too bulky residue and thus the shape of the entire molecule. Hence, it is essential to scale proton donor and acceptor properties in order to have a successive tool for molecular variations.

The concept of scaling solvent properties was comprehensively carried out by Kamlet and Taft[128], [129] who introduced the solvatochromic methology which correlates the UV wavelength shift of a probe and a reference molecule relative to an inert standard solvent in a series of organic media. Other solvent scales are Reichardt's E_T scale[130], the Y scale of Grunwald and Winstein[131] and the acceptor numbers of Gutmann[132].

However, what is actually described by the solvatochromic theory when the Taft value of an alcohol or amine is determined is the behaviour of OH or NH functions in the neat liquid. Here the OH or NH function is imbedded in a strong intermolecular hydrogen bond network which will enhance the hydrogen bond acidity of a single OH/NH function due to co-operative effects[133]. Those relations, however, may be completely different from those found when an OH or NH function interacts on a receptor group. Here 1:1 association would be the

appropriate model to describe the hydrogen bond acidity of the OH or NH function. This concept was taken up by Abraham et al.[134, 135, 136, 137, 138] who studied the equilibrium constants K of 1:1 proton donor-acceptor complexes in inert solvents and correlated logK with the hydrogen bond acidity and basicity of the associating molecules using the same proton acceptor and proton donor molecule, respectively. The equilibrium constants can be taken from the literature and were either received using optical spectroscopy under defined conditions (25 °C, low concentrations) or calorimetry.

By analysing the equilibrium constants K_A^H for a 1:1 hydrogen bond complexation of a series of acids and a reference base in carbon tetrachloride at 25 °C yield to a hydrogen bond acidity parameter α_2^H defined according to equation

$$\alpha_2^H = \frac{\log K_A^H + 1.1}{4.6363} \qquad [12]$$

Analogously a general hydrogen bond basicity parameter β_2^H was found which can be expressed by equation

$$\beta_2^H = \frac{\log K_B^H + 1.1}{4.6363} \qquad [13]$$

Here the 1:1 complexation constant for a series of bases and a reference acid was determined. Both equations can be combined to an expression for $\log K_{1:1}$ values in carbon tetrachloride[139].

$$\log K_{1:1} = 7.354 \cdot \alpha_2^H \cdot \beta_2^H - 1.094 \qquad [14]$$

Analogously to the overall hydrogen bond parameters $\sum\alpha_2^H$ and $\sum\beta_2^H$ the indices of α_2^H and β_2^H refer to 1:1 hydrogen bond associates but the parameters will never be identical since the conditions of the measurements are completely different. However, it would be expected that $\sum\alpha_2^H$ and α_2^H, and $\sum\beta_2^H$ and β_2^H are correlated to some extend.

Equation 14 is a cornerstone in the correlation of complexation constants with hydrogen bond parameters because now the correlation is independent on reference bases or acids and can be applied more generally.

For a selection of compounds the magnitude of α_2^H and β_2^H should be illustrated (Table 26).

Table 26: α_2^H and β_2^H values for a selection of compounds[140]

Compound	α_2^H	β_2^H
Water	0.353	0.38
Methanol	0.367	0.41
Ethanol	0.328	0.44
1.1.1.3.3.3-Hexafluoropropanol	0.771	0.03
Phenol	0.596	0.22
N-Methyl acetamide	0.383	0.72
Acetone	-	0.497
Ethylacetate	-	0.446
DMSO	-	0.774
Pyridine	-	0.625

It should be emphasised here that due to the occurrence of the electron lone pairs at the oxygen and nitrogen atoms the first six compounds can also act as proton acceptor molecules if other proton donor molecules are offered for association. Furthermore, water is not a particularly strong hydrogen bond acid and only of comparable strength as methanol and ethanol. The strong hydrogen bond effect of bulk water which need to be used to model aqueous media can be traced back to the property of acidification of each H_2O molecules due to intermolecular hydrogen bonds. The properties of bulk water will be difficult to describe because they will strongly relate to temperature and pressure.

The drawback of eq. 14 is that it correlates data measured in an apolar solvent which will lead to difficulties for the investigations of our peptide molecules which are insoluble in this solvent with some rare exceptions. Hence the research was continued in order to find similar relationships in different solvent systems[141]. Particularly we were interested in an equation for methylene chloride because most of the tripeptide derivatives are sufficiently soluble in this solvent and the proton donor properties (α_2^H= 0.129) of this solvent should not interfere with that of the 1.1.1.3.3.3-hexafluoropropanol (α_2^H= 0.771) when the association behaviour is studied.

From the literature equilibrium constants of associates studied in methylene chloride at 25 °C were selected (see table A12). The logarithm of these data are plotted vs. $\alpha_2^H * \beta_2^H$ in Fig. 21.

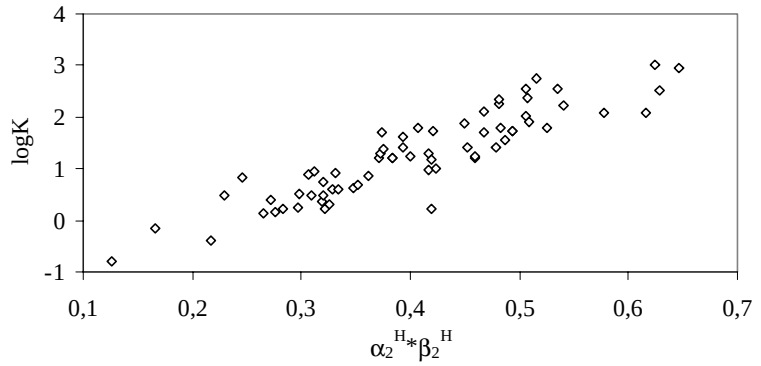


Fig. 21: Plot of $\log K_{1:1}$ vs. $\alpha_2^H * \beta_2^H$ of different 1:1 associate systems measured in methylene chloride

Although the linear regression would give a regression coefficient r^2 of 0.82 this value is not acceptable for our purpose. The reasons will be found in the specific interactions of the formed complex with the solutes which influences the result of the plot. However, if we distinguish between the different proton acceptor molecules we can enhance the correlation. Thus, just by selecting the acceptor compounds containing X=O, (X=C, S, P, As) and C-Y-C (Y=O, S) functions a correlation coefficient of 0.9807 (n=23) was calculated (Fig. 22).

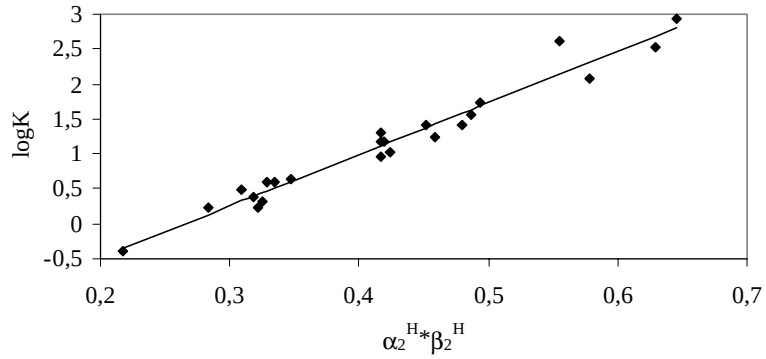


Fig. 22: Correlation of $\log K_{1:1}$ vs. $\alpha_2^H * \beta_2^H$ with acceptor compounds containing X=O, and C-Y-C functions (X=C, S, P, As, Y=O, S) measured in methylene chloride

From the linear regression analysis the coefficients of an equation can be set up which correlates the hydrogen bond acidity and basicity of esters, amides, sulfoxides and phosphates with alcohols in methylene chloride.

$$\log K_{1:1} = 7.3715 \cdot \alpha_2^H \cdot \beta_2^H - 1.9554 \qquad (n=23,\ r^2 = 0.9807) \qquad [15]$$

When studying peptide derivatives the problem of conformational interconversions always arises. The accessibility of an acceptor function will influence the association of particular functional groups within the molecule and thus the overall equilibrium constant as well. Before considering intermolecular association of the title compounds investigations of the conformational behaviour seems essential.

B.3.2 Conformational investigation on amino acid and peptide derivatives

The identification of protein and peptide structures by spectroscopic methods is well established and became an essential tool in the characterisation of this class of pharmaceutical compounds in the last decades.

Generally the protein conformation in crystalline probes which can be analysed by X-ray diffraction is strongly determined by the interplay of intramolecular and crystal-lattice forces and there are only a few examples in the literature where the intramolecular interactions still dominate[142]. The bioactive conformation of proteins is more likely found in solutions and especially aqueous media seems to match the biological conditions. Hence other spectroscopic methods are more applicable. The distinction of the method applied lies in its time scale and the effect which want to be studied. Hence conformational interconversions not requiring the rotation of the peptide bond may occur at milliseconds or faster time scale in linear peptides whereas the peptide rotations occur on the time scale of seconds[143]. NMR spectroscopy works on a slow time scale of seconds to hundreds of seconds. Therefore, it will give only averaged NMR parameters for conformational interconversions not requiring the peptide bond. This has to be kept in mind when analysing the NMR parameters. The advantage of this method lies in the individual observation of residues which can be analysed specifically if the spectrum is well resolved and the signals are assigned unambiguously.

Optical spectroscopy works at a faster time scale. Thus, excitation in the UV/VIS region of the electromagnetic spectrum takes place within 10^{-15} s and in the IR region within 10^{-13} s. Hence those methods are sensitive to conformation changes and allows the observation of different conformers simultaneously.

Especially vibrational spectroscopy offers a great potential for the investigation of peptide conformations as a complementary approach to NMR and crystallography. In cases of linear oligopeptide investigations it might be even the method of choice because the molecules are

not conformational constraint and hence NMR gives only averaged data. Unfortunately our samples do not crystallise and crystallography is not applicable.

The choice of an appropriate solvent for the spectroscopic investigation should be done with great care because it should not only fulfil the requirement of IR transparency and sufficient solubility of the oligopeptide derivatives but also represent the properties of a biological membrane which is a complex arrangement of zones of different polarity. Despite in aqueous phases the bioactive conformation in the apolar region of the lipid membrane are the subject of studies whereas the later environment is often simulated by apolar solvents. Mehler and Eichele[144] have estimated the dielectric properties of components of a lipid membrane in terms of the dielectric constants ε. The authors estimated an effective dielectric constants of 10-40 for the water accessible surface region of some proteins. At the hydrophobic side of the hydrocarbon bilayers of the phospholipids exhibits a dielectric constant ε of 2-5. Due to the specifics of the interactions which will take place when a molecule migrates through the phospholipid bilayer the truth of the dielectric constant will lie in an intermediate range in ε of 5-10. Table 27 gives dielectric constants of some solvents typically used in spectroscopy.

Table 27: Dielectric constant ε of selected solvents[145]

Solvent	Dielectric constant ε[a]
Water	78.30
Methanol	32.66
Acetonitrile	35.94
DMSO	46.45
1.1.1.3.3.3-Hexafluoropropanol[146]	16.70
1.2-Dichloroethane	10.37
Methylene chloride	8.93
Chloroform	4.81
Carbon tetrachloride	2.23
n-Heptane	1.92
n-Hexane	1.88
Vacuum	1

[a]relative permittivity of the pure liquid at 25 °C

Consequently the chlorinated methanes would be appropriate solvents for IR investigations. Among them chloroform exhibits a hydrogen bond acidity α_2^H of 0.197 whereas α_2^H of methylene chloride is 0.0 and thus a preferred solvent for association studies.

For the characterisation of the conformation different solvents should be tested which can be divided into four categories:

a) aqueous solvent
b) organic solvents with hydrogen bond donor abilities
c) organic solvents with hydrogen bond acceptor abilities
d) non-hydrogen bonding organic solvent

As representatives we have chosen water (H_2O or D_2O), 1.1.1.3.3.3-hexafluoropropanol (HFiP), dimethylsulfoxide (DMSO and DMSO-d6), and methylene chloride belonging to the categories a), b), c) and d), respectively.

The conformational behaviour of protected amino acids obeys other rules than that of peptide derivatives. Therefore, we have subdivided the following discussion accordingly.

B.3.2.1 Conformational behaviour of amino acid derivatives in different solutions

The amide groups of polypeptides and proteins possess nine characteristic vibrational modes or group frequencies[147, 148]. For compounds with amide groups Bellamy and Brügel[149] have generally classified these typical signals as amide I (1700-1670 cm^{-1}), amide II (1550-1510 cm^{-1}), amide III (ca. 1290 cm^{-1}), amide V (ca. 720 cm^{-1}, broad signal), amide IV and amide VI (ca. 600 and 620 cm^{-1}). Furthermore, the NH frequencies of amide groups are influenced by their surroundings and are assigned as amide A and B modes. Whereas the amide III, IV, V and VI bands are usually of low diagnostic value because they are found in the region of deformation vibrations in the infrared spectra which is in many cases overlapped by the solvent vibrations the amide I and amide II signals are well observable. The origin of these amide vibration and their sensitivity on conformational changes is determined by the coupling of typical carbonyl modes with that of the NH function. The keto-enol character of the amide function supports the coupling of the different modes. Furthermore, the keto-enol tautomerism of the amide group stabilised defined conformation and lowers the conformational flexibility of the peptide backbones. If amide I and II should be distinguished the amide I is originated to a large amount from the carbonyl stretching mode coupled with the NC mode of this functional group. In fact it can be classified as an antisymmetric OCN vibration. As Bellamy and Brügel explained the amide II originates from the coupling of the symmetric OCN vibration and the NH in-plane deformation mode. It is sensitive to any changes of the bond angle involving the NH function.

If using D_2O as an adequate substitute for H_2O the overlapping of the amide I and II signals by the deformation vibration of H_2O can be suppressed and enables the analysis of this spectral region even at lower concentrations and higher cell thickness. The hydrogen bond properties of H_2O are slightly different from that of its deuteriated oxide and small shifts of the signals due to this effect might be considered[150]. For Ac-Ala-OMe dissolved in the solvents belonging to the four categories the ester carbonyl and the amide I vibrations are shown in Fig. 23.

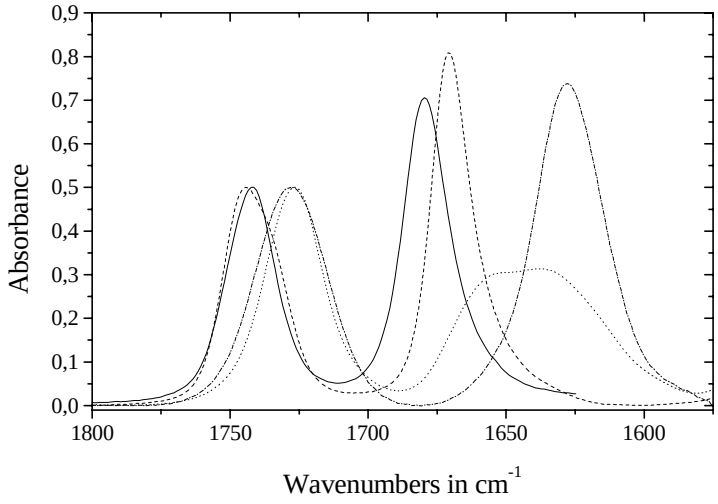


Fig. 23: Normalised spectra of the ester carbonyl and amide I signals of Ac-Ala-OMe dissolved in different solvents (solid line: CH_2Cl_2 dashed: DMSO-d6, dotted: HFiP and dot-dashed: D_2O)

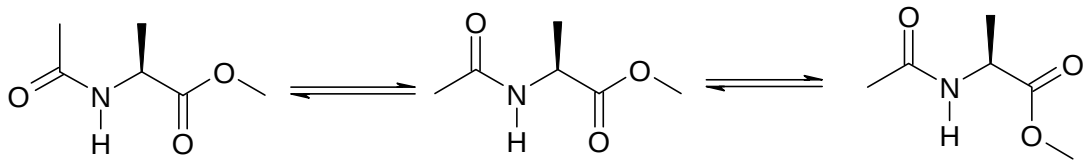
The observations of the solvent effect on the ester carbonyl signal makes it obvious that D_2O and HFiP act as deuteron/proton donors and due to the formed hydrogen bonds to the CO function the signals are shifted by ca. 15 cm^{-1} to lower wavenumbers. The proton donor capability of methylene chloride is rather weak and do not need to be considered here. The polarity of the solvent has only minor effects on the position of the absorption signals. DMSO-d6 exhibiting five times larger dielectric constant than CH_2Cl_2 causes almost no shift of the ester signal. Contrary it seems that the ester carbonyl band increases in the bandwidth and becomes asymmetrical. The reasons for this effect will be discussed later in detail.

The shift $\Delta\nu$ of the amide I signal in CH_2Cl_2 and DMSO-d6 can be numbered as 10 cm^{-1}. Due to the proton acceptor capabilities of DMSO-d6 the NH function is involved in a hydrogen bond and this reflects on the OCN vibration mode amide I. The amide I signal is not

symmetrical and the application of band deconvolution techniques will show that the profile can be fitted sufficiently using two Gauss-Lorentzian sum functions.

Analogously the effect of HFiP and D_2O on the carbonyl stretching signal reflects the proton donor properties of both solvents. In the amide I region, however, the fluorinated alcohol causes two signals. None of them is found at the position of the water solutions and hence we do not believe that the presence of water traces which cannot be completely avoided due to the hygroscopic properties of this solvent is the reason for the appearance of the band.

Some theoretical discussion about possible conformers should be given precedence over the more detailed information on solvent effect depending from the type of the amino acid shown in tables 28 to 31.



cis-amide **trans-amide**

Fig. 24: *Cis-trans* isomerism in Ac-Ala-OMe (For simplicity the protons of CH groups are not shown.)

Due to the rotation of the amide bond two different isomers might occur in the different solvents. Similar isomerism might also be found in the case of the ester function. The electron lone pairs at the ether oxygen are in conjugation with the carbonyl function. This will result in the stabilisation of a planar arrangement of this function and if steric repulsion of the side chain at C^{α} do not interfere might allow the internal rotation in this bond. The expected effects are, however, smaller than for the amide bond and it is, therefore, questionable whether the signals corresponding to both species can be separated. In the sketch of Fig. 24 it is also indicated that the NH group of the amide function is in-plane arrangement with the ester carbonyl function. Hence intramolecular interaction between these proton donor and acceptor functions should not be generally excluded from consideration although intramolecular hydrogen bonds forming five membered rings C_5 are not necessarily needed for the stabilisation of this geometry. Nevertheless, intramolecular hydrogen bonds of this kind are found in apolar solvents for a number of molecules. The extended conformation of peptides and proteins is comparable with the in plane arrangement of the amino acid unit. However, C_5

intramolecular hydrogen bonds are mostly not observed since proton donor solvent molecules interact with the acceptor functions.

As a model for description of amide functions in peptides and proteins by means of quantum mechanical calculations the small molecule N-methyl acetamide is often discussed in the literature. We have examined this molecules as well and added the band positions observed in the solutions in the following tables.

B.3.2.1.1 D_2O solutions

For the amino acid derivatives the positions of the band maxima should refer to the basicity of the functional groups in the aqueous environment (Table 28).

Table 28: Positions of the CO stretching and amide I bands of Ac-Xaa-OMe dissolved in D_2O [in cm^{-1}]

Compound	ν_{Ester}	$\nu_{Amide\ I}$
Ac-Ala-OMe	1728.0	1627.2
Ac-Val-OMe	1724.8	1628.3
Ac-Phe-OMe	1728.9	1630.7
N-Methyl acetamide		1623.0

In table 12 the effective hydrogen bond basicity of Ac-Ala-OMe, Ac-Val-OMe and Ac-Phe-OMe were estimated as equal to 1.10, 0.91 and 0.62, respectively. If the position of the ester carbonyl and amide I signal are taken as a first indication directly related to the basicity of these functional groups the corresponding signals should be shifted to lower wavenumbers with increasing basicity. The expected ranking of the positions was indeed found for the amide I signal but not for the ester carbonyl signal. The ester carbonyl signal seems to be influenced by the character of the amino acid side group. In the case of a branched residue on the C^{α} the ester carbonyl function was shifted to lower frequencies. However, in these small molecules also the mass effect might play a significant role since the increase of the mass will cause the shift of the signal as it would be expected from the theory. In larger molecules the mass effect can be neglected since its relative increase by changes of a peptide side function will be generally low.

B.3.2.1.2 HFiP solutions

Fluorinated alcohols are known to promote α- helix formation in peptide segments which are known to assume a β- structure under native conditions[151, 152]. Especially HFiP is one of the most potent alcohols known for inducing structures in proteins[153, 154]. However, there is still a controversy how the fluorinated alcohol influence the protein and peptide structure. The arguments in the literature is mainly focused on whether the direct interaction of the alcohol by the formation of hydrogen bonds causes the changes or if only the hydrophobic properties of the solvent result in the change of the secondary structure.

If considered on a molecular level HFiP is a stronger proton donor than water as the α_2^H indicates (Table 26) but a much lower proton acceptor. Consequently, HFiP will not form a hydrogen bond network similar to that known for water. This explains also why the dielectric constant of HFiP is lower than for water. Both phenomena reflect on the band positions of the amino acid derivatives. The ester carbonyl vibration is observed at almost identical frequencies than in D_2O: the ester carbonyl function is well exposed to the solvent molecules and enables the attack of the proton donor. Contrary, the amide I signals are not found at equivalent positions as in D_2O solution (Table 29). Assuming that in water solutions the D_2O molecules act as proton donors as well as proton acceptors when interacting with the ester carbonyl and the CO amide function and with the NH group of the amide, respectively, the extended conformation (middle) shown in Fig. 24 can be preserved.

HFiP exhibits only proton donor properties. Furthermore, it is a secondary alcohol with bulky CF_3 functions of high internal dipole moment. The energy minimisation of the Ac-Xaa-OMe -alcohol assemblies reveals that the alcohol induces a twist in the planar arrangement of the NH and the ester carbonyl function. In our case the change of the torsion angles cause two major conformers in the amide function which are less favoured than the hydrogen bond stabilised conformers in D_2O solution.

Table 29: Positions of the CO stretching and amide I bands of Ac-Xaa-OMe dissolved in HFiP [in cm^{-1}]

Compound	ν_{Ester}	$\nu_{Amide\ I}$	
Ac-Ala-OMe	1726.8	1660.9	1635.1
Ac-Val-OMe	1723.4	1662.7	1636.8
Ac-Phe-OMe	1727.5	1661.9	1637.8
N-Methyl acetamide			1624.4

The occurrence of the amide I signals for N-methyl acetamide at 1624.4 and 1623.0 cm^{-1}, for HFiP and D_2O, respectively, confirm our expectations regarding the strength of the hydrogen bond acidity of the fluorinated alcohol and water.

The amide II vibration of N-methyl acetamide is shifted strongly when using HFiP and D_2O as a solvent. In the amide II region two signals were found at 1514.1 cm^{-1} and 1493.1 cm^{-1} in D_2O solutions. From the resonance raman spectra[155] of the solutions and *ab initio* calculations[156] the signal at 1514.1 cm^{-1} can been assigned to the Fermi resonance between the amide II and a CH_3 deformation mode. In HFiP the amide II appears at 1572.1 cm^{-1}. This band is due to its origin sensitive to changes of the CNH bond angles. There are several explanations for this effect why the complex oscillating system contributing to the amide II group is changed. Firstly, due to the association at the CO amide group the electron density in the CN bond is changed in a different way than with the interaction with D_2O. Secondly, it might be a problem of the changed geometry caused by the bulkiness of the donor molecule. Consequently the dihedral angles could be different. Also the *cis-trans* ratio amide bond need to be considered. Williams[156] has performed quantum mechanical calculation on hydrated *cis* and *trans* N-methyl acetamide (NMA). It was shown that hydrogen bonding on *cis* or *trans* results in comparable stabilisation energies. *Cis/trans* isomerism in amide bonds of amino acid derivatives was confirmed by NMR spectroscopy if certain requirements are fulfilled. Hence the amino acid proline urges the formation of *cis* conformers because of the absence of a proton on the nitrogen atom[157, 158]. Furthermore, *cis* conformers are found in crystals and solution of Boc protected amino acids[159]. Obviously the acidic function stabilises the *cis* arrangement. VCD studies of N- urethanyl-L-amino acids of alanine, proline, valine and phenylalanine proved that in DMSO and chloroform solutions also *cis* isomers can be expected[160]. The amount of *cis* amide of Boc-Ala-OH was 85% and 94 %, in $CDCl_3$ and DMSO-d6 solutions, respectively. Summarising the information in the literature *cis/trans* isomerism is strongly dependent on the possibility of the formation of intramolecular hydrogen bonds which is favoured by the unprotected acidic function stabilising the *cis* form. With increasing polarity of the solvent and temperature the concentration of *cis* isomers increases. However, from our own ^1H NMR investigations of Ac-Ala-OMe in DMSO-d6 the integral areas underneath both $C^{\alpha}H$ signals reveal less than 2% of the *cis* form at 25 °C. In the spectra of Ac-Val-OMe and Ac-Phe-OMe no *cis* isomer could be detected. Also ^1H NMR spectra of the DMSO-d6 and D_2O solution give no evidence on higher concentrations of *cis*-NMA.

Based on molecular dynamic simulations Dwyer[161] has shown that in peptides the induced twist of the peptide backbone is a function of different factors (size, halogenation of the alcohol molecule, the hydrogen bond geometry, concentration and dielectric constant). In particular the [author] explained the helicogenic power of HFiP. It seems obvious that this effect is not restricted to peptides but affects already the conformers in protected amino acids.

B.3.2.1.3 DMSO-d6 solutions

DMSO is a typical proton acceptor agent and as such known to disrupt the secondary structure of globular proteins and polypeptides by destroying or weakening of the intramolecular hydrogen bonds [162]. In linear peptides it decreases strongly the population of β-turns due to the formation of solvent-peptide hydrogen bonds[163, 164]. Furthermore, DMSO has similar to HFiP bulky side chains. Hence when interacting with the proton donor function the torsion angles and consequently the band position of the amide signal might change (Table 30).

Table 30: Positions of the CO stretching, amide I and amide II bands of Ac-Xaa-OMe dissolved in DMSO-d6 [in cm^{-1}]

Compound	ν_{Ester}		$\nu_{Amide\ I}$	$\nu_{Amide\ II}$
Ac-Ala-OMe	1746.0 (52)	1733.8 (48)	1671.1	1545.4
Ac-Val-OMe	1741.9 (48)	1734.8 (52)	1671.1	1539.1
Ac-Phe-OMe	1744.5 (62)	1734.9 (38)	1672.1	1545.7
N-Methyl acetamide			1667.7	1556.6

The application of band separation methods indicates the overlapping of two bands in the ester carbonyl region. The numbers in the brackets give the percentage of integral amount of the conformer. Almost to equal amounts the two conformers at the ester group are found which causes the asymmetric band shape of the ester carbonyl signal. Due to the interaction between DMSO and the NH function changes the conformer distribution at the C- terminal protecting group. In case of Ac-Phe-OMe the side chain seems to influence the conformational equilibrium.

The slight asymmetric band shape of the amide I mode results from the presence of water traces which unfortunately cannot be avoided during sample preparation.

B.3.2.1.4 Methylene chloride solutions

This solvent does not reveal as a proton donor and acceptor. Hence besides dielectric interactions the molecules should behave almost undisturbed. Fig. 25 pictures the spectral region of the acetyl amino acid methyl esters dissolved in methylene chloride. In comparison to the spectra recorded from the D_2O solutions the ester and the acetyl amide function at ca. 1742 cm^{-1} and 1680 cm^{-1} (Ac-Ala-OMe), respectively, are shifted to higher wavenumbers. But also in this solvent the influence of the type of the amino acid is obvious.

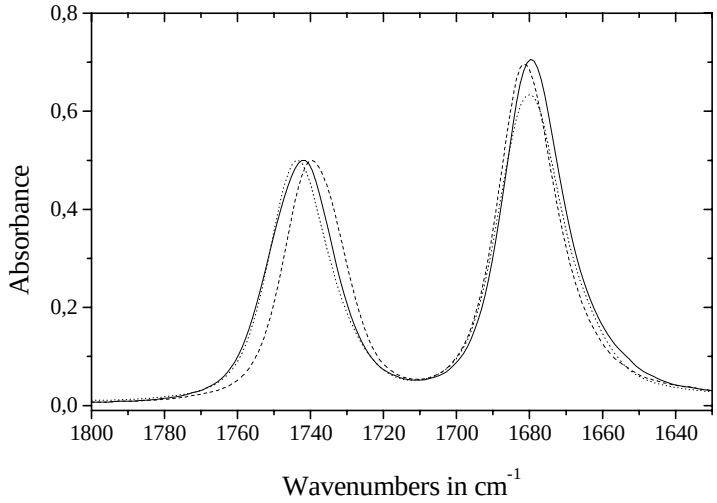


Fig. 25: Normalised spectra of the ester and amide I vibration modes of Ac-Xaa-OMe dissolved in methylene chloride (solid line: Ac-Ala-OMe, dashed: Ac-Val-OMe, dotted: Ac-Phe-OMe,

In a homologous series of compounds it would have been expected that the position of the band maxima follows the rules of the mass effect. Hence with increasing mass significant signals should relatively shift to lower wavenumbers due to the change of the oscillator properties if no differences in the population of conformers occurs. Interestingly the compound with the branched side chain at C^{β}, Ac-Val-OMe does show an ester carbonyl signal at lowest wavenumbers in methylene chloride (1739.5 cm^{-1}) and D_2O (1724.7 cm^{-1}) whereas the amide I signal shows the inverse effect: in methylene chloride the position of the amide bands shifts to lower wavenumbers in the order Val >Ala ≥ Phe; in D_2O the band position follow the order Phe >Ala >Val. Table 31 confirms the influence of the C^{α} substitution also for Z- protected amino acid esters where the urethane band was observed at 1724.0 cm^{-1}.

Table 31: Positions of the carbonyl and amide I signals of the amino acid derivatives in CH_2Cl_2 after band fitting [in cm^{-1}]

Compound	ν_{Ester}	$\nu_{Urethane}$	$\nu_{Amide\ I}$
Ac-Ala-OMe	1742.4		1679.9
Ac-Val-OMe	1739.0		1681.6
Ac-Phe-OMe	1743.1		1680.3
Boc-Val-OMe	1740.9	1714.0	
Z-Val-OMe	1742.2	1724.0	
Z-Phe-OMe	1746.0	1723.9	
N-Methyl acetamide			1674.7

B.3.2.1.4.1 NH band profile

In less polar media also the NH stretching vibration can contribute to the discussion about conformational equilibria and even more important to the occurrence of intramolecular hydrogen bonds which would favour specific arrangements. The NH stretching region of the valyl derivatives dissolved in methylene chloride is shown in Fig. 26.

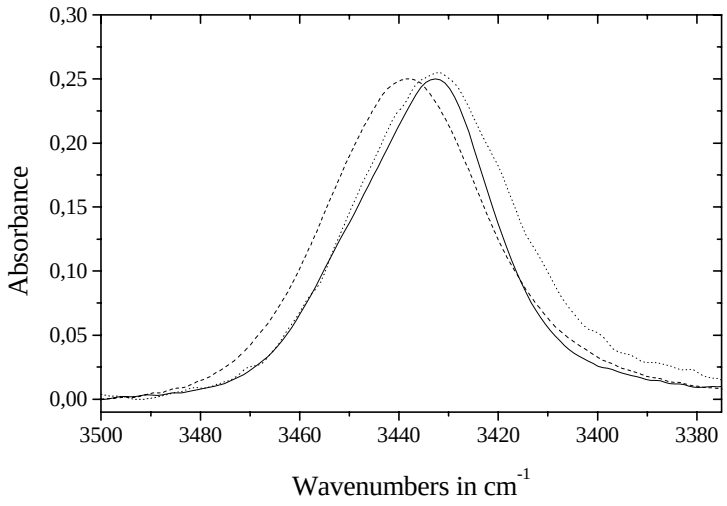


Fig. 26: Normalised NH band profile of valyl derivatives with different N-terminal protecting groups dissolved in CH_2Cl_2 (solid: Ac-Val-OMe, dashed: Boc-Val-OMe; dotted: Z-Val-OMe)

The asymmetric band profile can be fitted using Voigt functions. The positions of the separated signals and the ratio of the integral intensities are given in table 32.

Table 32: NH band positions of Ac-Xaa-OMe dissolved in CH_2Cl_2 after band fitting [in cm^{-1}]

Substance	ν_{NH1}	ν_{NH2}	Intensity ratio NH2/NH1
Ac-Ala-OMe	3447.1	3432.5	5.72
Ac-Val-OMe	3448.4	3431.3	4.11
Ac-Phe-OMe	3440.6	3427.2	4.53

Similar behaviour of an asymmetric NH band was observed in carbon tetrachloride for Z- and Boc protected amino acid esters. Besides overtone and Fermi resonance signals the occurrence of *cis/trans* isomerism and intramolecular hydrogen bonding can be named as possible reasons for the asymmetrical band shape. From the arguments given above *cis/trans* isomerism can be excluded. Former near infrared studies and H/D exchange experiments performed on Z- protected amino acid derivatives have shown that overtone or Fermi resonance signals do not overlap with the NH stretching signal[165]. Hence we investigated the NH overtone signal of NMA and Ac-Val-OMe in methylene chloride solutions. Fig. 27 shows that the overtone signal of Ac-Val-OMe is asymmetrical whereas the band shape of NMA is highly symmetrical. Conclusively the Fermi resonance does not occur. We have to consider intramolecular hydrogen bonding forming a C_5 ring for an explanation of the effect.

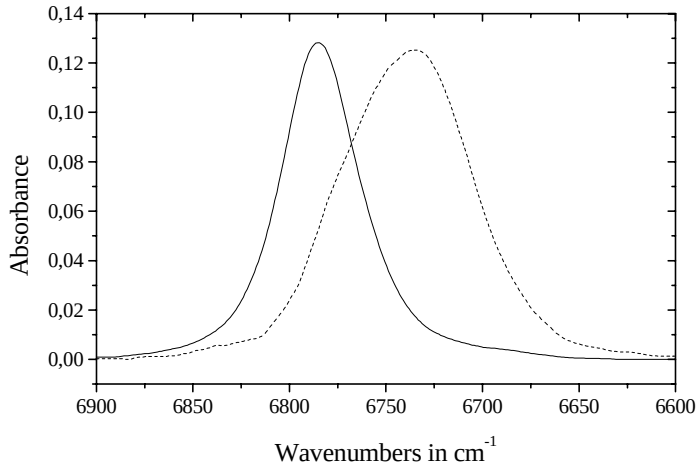


Fig. 27: Normalised NH overtone signals of NMA (solid) and Ac-Val-OMe (dashed) dissolved in methylene chloride

For Boc-Val-OMe, Z-Val-OMe and Z-Phe-OMe an asymmetric NH band profile was observed in n-hexane, cyclohexane and carbon tetrachloride solutions. Unambiguously the signal at low wavenumbers was assigned to the C_5 intramolecular hydrogen bonded associate. Using similar band fitting procedures band maxima at 3465.0 and 3443.5 cm^{-1}, 3460.2 and

3438.6 cm^{-1}, 3454.6 and 3435.1 cm^{-1}, exhibiting a intensity ratio of 4.34, 4.17 and 4.69 in the carbon tetrachloride solutions of Boc-Val-OMe, Z-Val-OMe and Z-Phe-OMe, respectively, could be identified. Based in an iterative approach recognising the temperature dependence of the extinction coefficient the equilibrium constants could be approximated as equal to 4.6-2.3 and 3.2-2.0, for Z-Val-OMe and Z-Phe-OMe, respectively[166]. The enthalpy for intramolecular C_5 association was calculated as ca. -1 kJ/mol. However, in methylene chloride solution the corresponding NH signals of Z- and Boc- protected amino acid methyl esters were symmetrical (Fig. 26). Considering the solvent effect the two NH signals observed in methylene chloride solutions of Ac-Ala-OMe, Ac-Val-OMe and Ac-Phe-OMe have to be assigned to a free NH (at ca. 3445 cm^{-1}) and a C_5 associated NH (ca. 3430 cm^{-1}) group. Maxfield et al.[167] came to similar conclusions in their investigations on N- methyl acetamides in $CHCl_3$ solutions. It was rather surprising to find C_5 associates in CH_2Cl_2 but obviously for the stability of C_5 intramolecular hydrogen bonds the kind of the N- terminal protecting group matters; whereas NH functions linked to urethane groups form C_5 associates only in carbon tetrachloride N- acetyl groups stabilises this arrangement even in chloroform or methylene chloride.

B.3.2.2 Conformational behaviour in oligopeptide derivatives

With enlarging the numbers of amino acid units the conformational variety increases drastically until the formation of intramolecular hydrogen bonds and the coupling between transition dipoles stabilises certain secondary structure. In polypeptides or proteins these helical and β-sheet arrangements give typical infrared spectra. In particular, the amide I and II signals can be used as good indicators for monitoring the secondary structures in larger peptides and proteins. As an example, poly-L-lysine adopts particularly a α-helical conformation giving an amide I band at 1638 cm^{-1}. The β-sheet band frequency can be located at 1610 cm^{-1} whereas the weaker band at 1680 cm^{-1} is associated with the high frequency vibration of antiparallel β- sheet structure[168, 169].

Based on second derivative and deconvolution techniques the analysis of bandwidths and intensities of overlapping bands is possible[170, 171] and allows the quantitative identification of secondary structures present in peptide molecules under the assumption that the effective intrinsic absorptivities of the amide I bands corresponding to different structures are very similar [172, 173].

Characteristic features of oligopeptides are so called turns which also exist in polypeptides and define a site where the chain changes its overall direction. Two major classes of turns may be specified according to the size of the loop which may or may not be stabilised by intramolecular hydrogen bonds. In β-turns the C=O residue *i* may interact with the NH of the residue *i+3* while in γ-turns the C=O of residue *i* may be hydrogen bonded to the NH residue *i+2*. In terms of the size of the formed intramolecular hydrogen bond rings β- turns will result in 10-membered associate rings, C_{10}, which would be equal to a loop of the 3_{10} helix. Contrary, the number of members involved in an associate ring of γ- turns is seven. If intramolecular hydrogen bonds are formed we will follow the nomenclature of Bragg et al.[174] to characterise the size of the associate rings, C_{10} and C_7, throughout the text. Analogously C_5 will indicate a five membered associate ring which can occur by the interaction of the NH and the CO function of the same amino acid residue which is, however, rarely found in aqueous media because the solvent interactions are stronger than the intramolecular hydrogen bonds. Based on the dihedral angles the C_5 arrangement will lead to an extended peptide chain conformation.

Intramolecular hydrogen bonding in β- and γ-turns requires defined dihedral angles φ and ψ of the peptide backbone. In γ-turns the dihedral angles are not strictly defined and can vary between 70 to 85 and -60 to -70 for ϕ and ψ, respectively. In inverse γ-turns the angles ϕ and ψ have the same magnitude but with the opposite signs. Fig. 28 depicts both γ-turns.

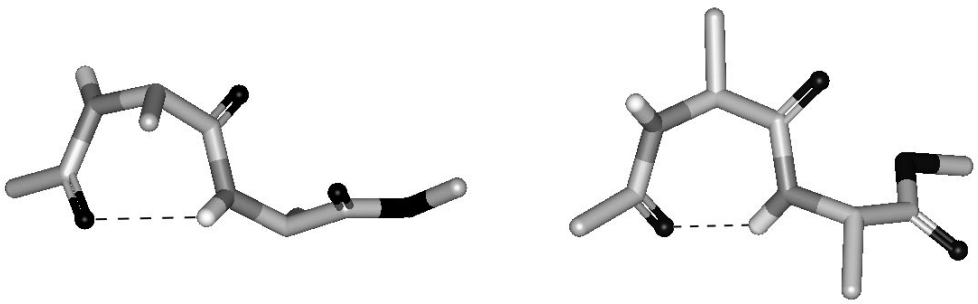


Fig. 28: γ-turn (left) and inverse γ-turn (right) sketched for Ac-Ala-Ala-OMe (colour code: black: oxygen, grey: nitrogen, light grey: carbon, white: hydrogen in NH; otherwise, the hydrogen atoms are not shown)

The strength of the C_7 intramolecular hydrogen bond should be affected if the side chain residue interferes. Hence in terms of C_7 hydrogen bond rings an axial and equatorial form can be distinguished. The later is more abundant in proteins and corresponds to an inverse turn[175].

β-turns which may be stabilised by ten membered intramolecular hydrogen bond rings can be divided into different types -type I (I'), type II (II') type III (III')- which are sketched in Fig. 29[107]. Type IVa and b are not shown because they are only relevant for *cis*-amide bonds which are generally found in proline derivatives. It should be noted that for Ac-Ala-Ala-Ala-OMe an other C_{10} associate would be possible involving the ester carbonyl function and the NH of ^{2}Ala.

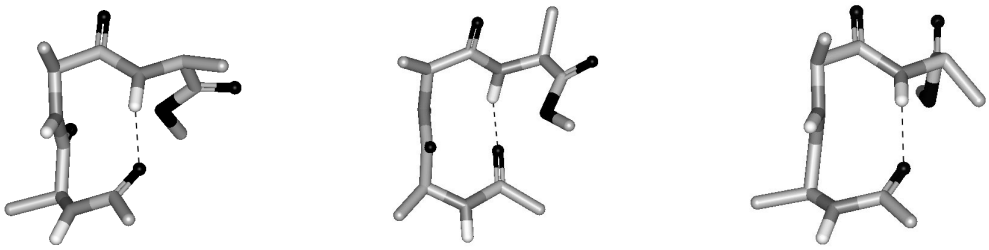


Fig. 29: β- turns of type I, type II and type III (from left to right.) which can theoretically be formed in peptides with more than three peptide bonds. Here it is depicted on the example of Ac-Ala-Ala-Ala-OMe. (colour code: black: oxygen, grey: nitrogen, light grey: carbon, white: hydrogen in NH; otherwise, the hydrogen atoms are not shown)

From theoretical calculations only a limited number of ϕ, ψ backbone dihedral angle combination would reveal for the i^{th} residue of a peptide[176, 177]. The solution state conformation of oligopeptides is often claimed to be random[178, 179] unless conformational constraints stabilise intramolecular interactions. Thus, γ-turns as depicted in Fig. 28 for the example of Ac-Ala-Ala-OMe have been frequently observed in the solid state in a slightly distorted form but in solution the dominance or singular presence of such a structure has been found only in a few cases in apolar solvents[180, 181, 182].

Also β- turns are sensitive to environmental effects. Hence 1→4 hydrogen bonds may break apart when dissolving the compound. Therefore, the literature on turn structures in oligopeptides in solutions refers to conformationally constrained molecules such as cylcopeptides[183, 184, 185]. Examples are described in the literature where also γ- turns exists in solutions of cyclopeptides[186, 187].

In linear peptides the intramolecular hydrogen bond often formed between acidic C- or N-terminal functions in polar media[164, 188, 189]. Only rare cases are reported on protected peptide molecules being stabilised by turn structures in hydrocarbon solvents[190].

Generally, linear oligopeptides are supposed to occupy a „random" orientation of the peptide backbone. Following the suggestion of Woody[191] we would rather use the term „unordered" throughout this work. We feel that „random" is not appropriate since it implies

that there are no preferred conformers. Contrary, although intramolecular stabilisation by hydrogen bonding does not occur the polar interactions of the solvent surrounding will support the occurrence in certain conformers depending on the solvent polarity and the structure of the compound. Our discussion of the FTIR data will be not focused on an unambiguous assignment of the signals but on relative changes in the population of conformers due to the variation of the amino acid residue, the peptide length, the solvent polarity and the temperature.

B.3.2.2.1 Effect of peptide chain length

The asymmetric band shape of the amide I signals of Ac-Ala-OMe, Ac-Ala-Ala-OMe, Ac-Ala-Ala-Ala-OMe and Ac-Ala-Ala-Ala-Ala-OMe dissolved in D_2O depicted in Fig. 30 indicate the heterogeneous distribution of conformers.

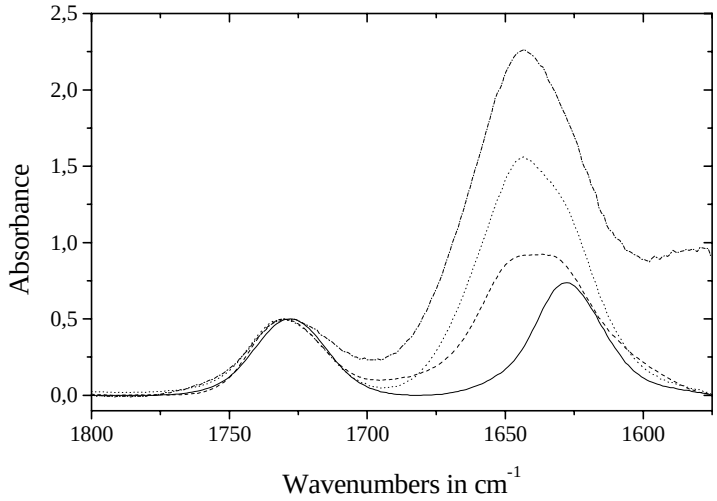


Fig. 30: Normalised spectra of the ester carbonyl stretching and amide I region of Ac-Ala-OMe (solid), Ac-Ala-Ala-OMe (dashed), Ac-Ala-Ala-Ala-OMe (dotted) and Ac-Ala-Ala-Ala-Ala-OMe (dott-dashed) dissolved in D_2O

From the positions of the band maxima the signal of the methyl ester at ca. 1730 cm^{-1} and the amide I absorption at 1625 and 1630 cm^{-1} can be assigned unambiguously. With the extension of the molecule by a second amino acid two amide I signals can be detected at higher wavenumbers than found for N-acetyl amino acid methyl esters. However, it is obvious that the character of the amide functions must have been changed compared to Ac-Ala-OMe by varying the dihedral angles.

It should be noted here that the enlarging of the dipeptide by another amino acid does not result in the occurrence of an additional amide band. In fact, the amide I signals of Ac-Ala-Ala-Ala-OMe and Ac-Ala-Ala-Ala-Ala-OMe show almost identical band profiles only with a higher intensity in the tetrapeptide.

B.3.2.2.2 Effect of solvent polarity

Although linear peptides are supposed to exist in a mixture of different conformers we believe that similar to constrained peptides certain conformers should be preferred. In the NH stretching signal (amide A) of the methylene chloride solutions the occurrence of intramolecular hydrogen bonds cannot be confirmed. Nevertheless, it is well accepted in the literature that hydrophobic interactions should be an important if not the major factor stabilising secondary structures. Hence if the properties of the side chain do restrict the flexibility among the variety of conformers some may be privileged.

In order to differentiate the existing conformers the band profile of the amide I (and II) were examined using band deconvolution and band fitting methods with a set of Gauss-Lorentzian sum functions. Fig. 31 exemplifies the curve-fitted spectrum for the DMSO-d6 solution of the tetrapeptide Ac-Ala-Ala-Ala-Ala-OMe.

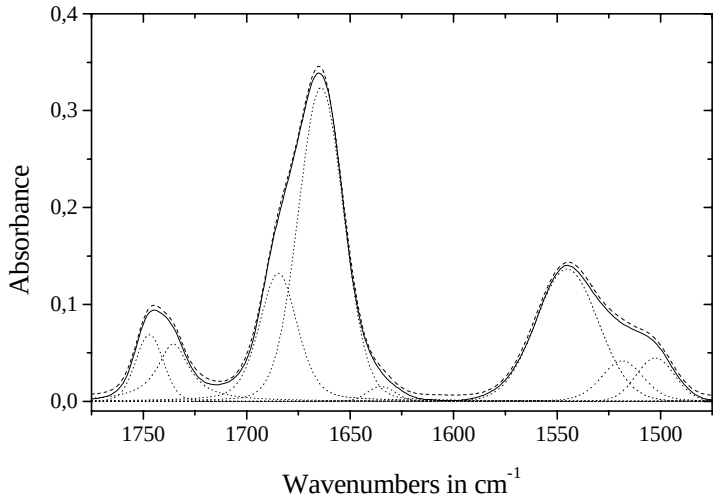


Fig. 31: Spectra of the ester and the amide I and amide II region of Ac-Ala-Ala-Ala-Ala-OMe dissolved in DMSO-d6 and the fitted components

The ester carbonyl signals can be assigned based on the knowledge from the protected amino acid derivatives. Note that with enhancing peptide chain only a small shift of the ester signal was found. Hence the mass effect is relatively small (Table 33).

Table 33: Positions of the ester carbonyl band of alanyl derivatives dissolved in different solvent after application of band deconvolution and band fitting methods [in cm^{-1}]

Compound	D_2O	HFiP	DMSO-d6	CH_2Cl_2	Assignment
Ac-Ala-OMe			1746.8		
	1728.0	1726.8	1735.3	1742.4	ester CO
Ac-Ala-Ala-OMe			1747.3	n. s.	
	1729.2	1728.7	1736.1		ester CO
Z-Ala-Ala-OMe	n. s.		1747		
		1728.0	1735.5	1742.3	ester CO
Ac-Ala-Ala-Ala-OMe			1747.5	n. s.	
	1729.9	1728.1	1736.1		ester CO
Z-Ala-Ala-Ala-OMe	n. s.		1747.4		
		1728.0	1735.9	1743.5	ester CO
Ac-Ala-Ala-Ala-Ala-OMe			1716.9	n. s.	
	1728.9	1728.5	1735.4		ester CO

n.s. = not soluble

The solvent effect on the amide I signal can be estimated from N-methyl acetamide in which the amide function should be fully accessible for the solvent attack. Taking the amide signal at 1674.7 cm^{-1} of N-methyl acetamide dissolved in methylene chloride as a reference the solvent induced shift $\Delta\nu$ can be numbered as -7, -50.4 and -51.6 cm^{-1} in DMSO-d6, HFiP and D_2O, respectively. This estimation legitimates us to assign the signals of amide functions which are fully accessible for the solvent. From theoretical consideration it may be assumed that this conformer must be similar to the extended conformation.

In table 34 we have distinguished between mono furcated and solvent exposed amide I species. The term „solvent exposed“ is used to characterise different types of interactions: When proton donors act as solvent strong intermolecular hydrogen bond will occur. When using DMSO-d6 or CH_2Cl_2 simply dipolar interactions will be dominant. Here conformational equilibria can also lead to more than one solvent exposed group. In terms of hydrogen bonding between the amide function and D_2O or HFiP the functions are completely bonded. Mono furcated species, however, only interact with one proton donor molecule and are, therefore, observed at higher wavenumbers. Furthermore, a generally small amount of intermolecularly hydrogen bonded species are assigned which result from homoassociation of the peptide molecules.

Table 34: Positions of the amide I signal of alanyl derivatives dissolved in different solvent after application of band deconvolution and band fitting methods [in cm^{-1}]

Compound	D_2O	HFiP	DMSO-d6	CH_2Cl_2	Assignment
Ac-Ala-OMe		1660.6 (18)			mono furcated
		1636.4 (70)	(64)	1679.9 (80)	solvent exposed amide
	1627.2 (100)	1614.8 (12)	1659.2 (36)	1667.1 (20)	intermol. ass. species
Ac-Ala-Ala-OMe	1650.5 (51)	1672.5 (22)		n. s.	mono furcated
	1630.1 (45)	1645.9 (61)	1683.0 (28)		solvent exposed
			1664.7 (65)		solvent exposed
	1601.6 (4)	1614.7 (17)	1646.6 (7)		intermol. ass. species
Z-Ala-Ala-OMe	n. s.	1701.8	1715.2	1720.4	urethane
		1670.2 (39)			mono furcated
		1645.4(61)	1681.5 (78)	1683.9 (100)	solvent exposed
			1666.4 (22)		solvent exposed
Ac-Ala-Ala-Ala-OMe	1646.1 (75)	1669.2 (51)		n. s.	mono furcated
	1624.8 (25)	1635.1 (49)	1683.7 (17)		solvent exposed
			1664.0 (81)		solvent exposed
			1635.1 (2)		intermol. ass. species
Z-Ala-Ala-Ala-OMe	n. s.	1700.6	1714.5	1720.1	urethane
		1669.0 (63)			mono furcated
		1642.8 (37)	1689.0 (27)	1688.5 (36)	solvent exposed
			1668.9 (69)	1671.3 (64)	solvent exposed
			1640.4 (4)		intermol. ass. species
Ac-Ala-Ala-Ala-Ala-OMe	1668.3 (4)	1669.0 (46)		n. s.	mono furcated
	1646.5 (59)	1641.7 (36)	1684.7 (27)		solvent exposed
	1626.1 (37)	1617.4 (18)	1664.0 (71)		solvent exposed
			1634.9 (2)		intermol. ass. species

n.s. = not soluble

In the literature the assignment of turn structures in cyclopeptides is often based on normal mode analysis. According to these calculations the characteristic band for different types of β-turns should be found at 1640 ±2 cm^{-1}[192]. However, recent results have shown that the distinction of the type of the β- turns based on the calculation of the hydrogen bond induced variation of the CO dipole moment may simplify the situation because also coupling between the transition dipoles may also lead to conformation dependent effects[193]. Hence the assignment of type I, II and III β- turns need to be re-thought in order to explain the discrepancies between theoretical prediction and experimental findings[194]. Furthermore, solvent effects on the position of the characteristic signals have been usually neglected.

Our interpretation of the band positions in linear peptides is, therefore, different from literature data. But we feel that a signal at 1640±2 cm^{-1} in our substances is coincidental at the same position as β- turn signals in cyclopeptides structure and should be rather an effect of solvent interaction and conformational variances in the peptide backbone.

Comparing the Z- and the Ac- protected peptides the occurrence of bands at similar positions can be found which will help for the assignment in case of the insolubility of the compound in a certain solvent. As a general trend it was found that with enlarging the molecule the solvent exposed amide function in methylene chloride and DMSO-d6 shifts to higher wavenumbers. Obviously this function becomes more shielded against the solvent attack. Similar observation can be done for aqueous media. With increasing chain length the amide functions are not solvated in the same manner. Hence, in Ac-Ala-Ala-OMe two bands of similar intensity are found at 1650.5 and 1630.1 cm^{-1}, the later one occurring at exactly the same positions as the amide I in Ac-Ala-OMe. A value for Δν of ca. 20 cm^{-1} is quite large to be explained by H/D exchange which will undoubtedly take place. One could argue supported from the intensity ratio of both bands that both signals correspond to the different peptide functions of the molecule. As the conclusion the question will arise why similar observations cannot be found in other solvents. Hence we tend to assume that the signals are an effect of the interaction of one or two D_2O molecules per amide group and causes a mono furcated and bi furcated associate species. The symmetry of association cannot be proved and hence we indicated the interaction with more than one donor molecule as solvent exposed.

In the brackets the relative intensities of the amide I signals are given. In the literature the intensity ratio is often used as a tool for the assignment[187] . We give these values to indicate a tendency but not for quantitative conclusions since the molar extinction coefficient is not known and it is proved by Krimm et al.[195] that the extinction changes drastically with the conformation.

B.3.2.2.3 Effect of amino acid substitution

We have systematically discussed the amide I signal homopeptide derivatives which were built from alanyl amino acid units. Now the effect of amino acid substitution will be shown on the hand of the methylene chloride solutions of selected tripeptide derivatives (Fig. 32). Naturally, with increasing number of amino acid residues the interpretation of the spectra becomes difficult.

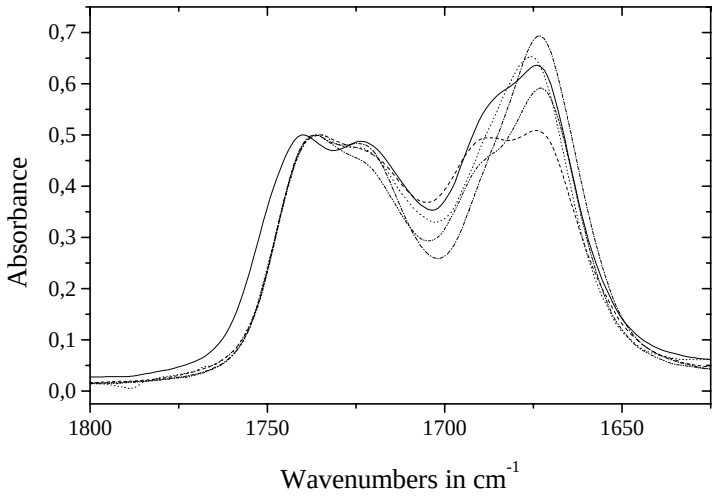


Fig. 32: Normalised spectra of the ester carbonyl, urethane and amide I region of tripeptide derivatives dissolved in CH_2Cl_2 (solid: Z-Ala-Ala-Ala-OMe, dashed: Z-Ala-Ala-Val-OMe, dotted: Z-Ala-Val-Val-OMe, dot-dashed: Z-Val-Val-Val-OMe and dot-dot-dashed: Z-Val-D-Val-Val-OMe)

Two major conclusions can be drawn. Firstly, the amide I profile indicates the overlapping of different conformers. The position of amide I band of each conformer is determined by the kind of the amino acid and its chirality. Secondly, the urethane function is changed in intensity and position due to the substitution. Again the observation substantiate that the amide I band shape is the result of different conformers rather than the number of peptide bonds. After band deconvolution the position and relative integral intensities (in brackets) of the separated amide I signals can be given (Table 35).

From the analysis of the band position and the intensities it becomes obvious that the kind of the amino acid residue, the sequence and the chirality determine the amide I band profile but the changes do not seem to be systematically. We have extended the band profile analysis to tripeptide derivatives Z-Ala-Xaa-Val-OMe and numerous diastereoisomers but the results do not give any further information.

Table 35: Band positions of dipeptide and tripeptide derivatives with amino acid substitution

Substance	ν_{Ester}	$\nu_{Urethane}$	$\nu_{Peptide\ I}$	$\nu_{Peptide\ II}$
Z-Ala-Val-OMe	1740.2	1720.6	1685.7	
Z-Val-Val-OMe	1739.6	1721.1	1682.7	
Z-Phe-Phe-OMe	1744.0	1722.9	1682.7	
Z-Ala-Ala-Val-OMe	1740.7	1722.7	1690.7 (36)	1671.7 (64)
Z-Ala-Val-Val-OMe	1741.2	1723.1	1688.7 (49)	1671.0 (51)
Z-Val-Val-Val-OMe	1741.2	1723.1	1688.1 (22)	1671.8 (78)
Z-Val-D-Val-Val-OMe	1741.2	1723.1	1691.2 (27)	1671.5 (73)
Z-Phe-Phe-Phe-OMe	1744.9	1722.4	1687.1 (38)	1670.6 (62)
Z-Ala-Val-Leu-OMe	1743.6	1721.5	1688.3 (29)	1674.0 (71)
Z-Ala-Leu-Val-OMe	1741.4	1723.5	1687.7 (39)	1675.0 (61)
Z-Ala-Phe-Val-OMe	1742.2	1724.9	1687.9 (49)	1671.0 (51)
Z-Ala-Val-Phe-OMe	1744.8	1721.9	1691.0 (19)	1674.3 (81)

The temperature behaviour of the amide I signals was tested for homodipeptides and homotripeptides in the temperature range between 5 and 30 °C. At the example of Z-Phe-Phe-Phe-OMe it is shown that with raising temperature both amide bands shift to higher wavenumbers while the intensity of the signals decreases due to the changes of the extinction coefficient (Fig. 33). Furthermore, it can be noted that the temperature effect of the ester carbonyl signal is ca. 0.02 cm^{-1}/K smaller than for the urethane and the peptide signals. Here temperature induced shifts of the band positions of 0.04 and 0.07 cm^{-1}/K were observed, respectively. We would interpret these findings as an expression for the conformational heterogeneity of molecules with highly flexible backbone.

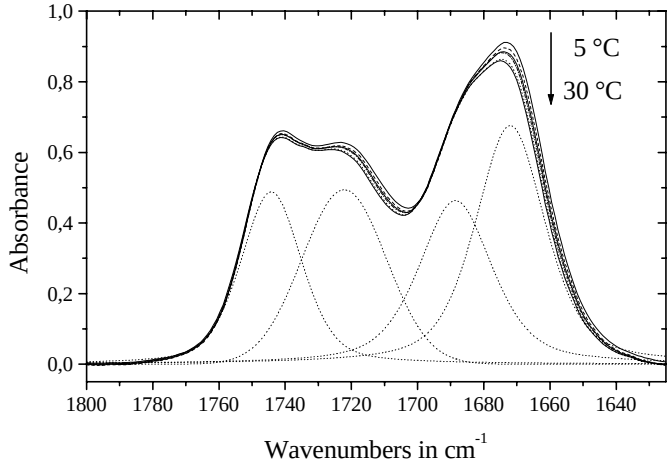


Fig. 33: Temperature effect on the signals between 1800 and 1625 cm^{-1} of Z-Phe-Phe-Phe-OMe dissolved in methylene chloride

B.3.2.2.4 CD measurements in different solvents

The NH band profiles of the investigated substances dissolved in methylene chloride do not show any associates which could be assigned to C_7 hydrogen bond rings. Thus, there is no stabilisation of γ- turns by intramolecular hydrogen bonds in this solvent.

In the assignment of the amide I signals we have claimed that the occurrence of a band at ca. 1640 cm^{-1} does not indicate a β- turn but is an effect of solvent interaction. To prove the correctness of the conclusions drawn from the infrared spectra CD measurements were performed since the different types of β- turns give characteristic patterns in the vacuum UV region[191]. Hence it was predicted that for the most common types expected in proteins the type I, II and III (the latter being a variant of type I and corresponding to one loop of a 3_{10} helix), the CD spectra should have a weak, red-shifted negative $n\rightarrow\pi^*$ band near 225 nm, a strong positive $\pi\rightarrow\pi^*$ band with a maximum between 200 and 205 nm, and a strong negative band between 180 and 190 nm. CD spectra containing these features are termed Class B. It can be distinguished from that of β- sheets (Class A) by a 5-10 nm red shift of the maxima. Type I (III) turns generally give a Class C CD pattern which is slightly blue-shifted and α-helix like spectra with decreases band intensities. Typical λ_{max} ($[\theta]_{MR}$) parameter of a C spectrum measured in TFE are 216 nm (-4.600), 205 nm (-6.200) and 186 nm (6.700)[196]

In unordered polypeptides or proteins the characteristic features in the chirograph are a strong negative band just below 200 nm and a weaker signal near 170 nm in the vacuum UV region[197].

In order to meet the requirements of vacuum-UV transparency we investigated Ac-Val-OMe, Z-Val-Val-OMe and the diastereoisomeric tripeptide derivatives Z-Val-Val-Val-OMe (LLL and LDL) in acetonitrile and HFiP in the region between 180 and 260 nm. The choice of acetonitrile as adequate solvents for methylene chloride seems reasonable although it cannot be neglected that weak proton acceptor interaction initiated by the solvent will occur.

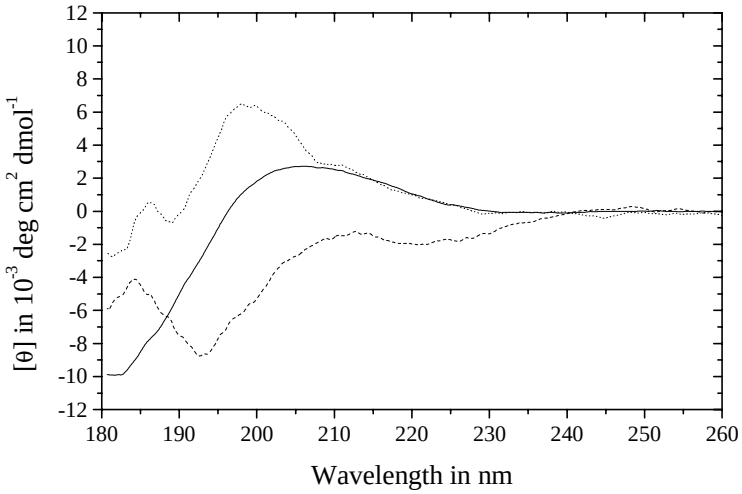


Fig. 34: CD spectra of Ac-Val-OMe (solid line) and the diastereoisomers of Z-Val-Val-Val-OMe (dashed: LLL, dotted: LDL) dissolved in HFiP

In general the patterns of the CD spectra in both solvents were similar and hence only the chirograph of the compounds dissolved in HFiP is shown in Fig. 34. As an effect of the solvent polarity the maximum of the CD band for Ac-Val-OMe is shifted from 214 nm in acetonitrile solutions to 206 nm in the proton donor solvents HFiP. Although the signal-to-noise ratio of the tripeptide chirographs is rather poor the main features are found in both solutions: the LDL diastereoisomer exhibits a strong positive signal at 199 nm which has a shoulder at higher wavelength localised at ca. 217 nm. Contrary, for the LLL diastereoisomer a intense negative CD band with a minimum at 197 nm and a weak band at 214 nm were observed. The intense signals in the tripeptide spectra are caused by $\pi \rightarrow \pi^*$ transitions of the aromatic ring of the benzoyloxycarbonyl protecting group. This function significantly determines the spectra and as the temperature dependent study of Z-Val-Val-Val-OMe shows a trough assigned to the $\pi \rightarrow \pi^*$ transition of the phenyl ring decreases with raising temperature.

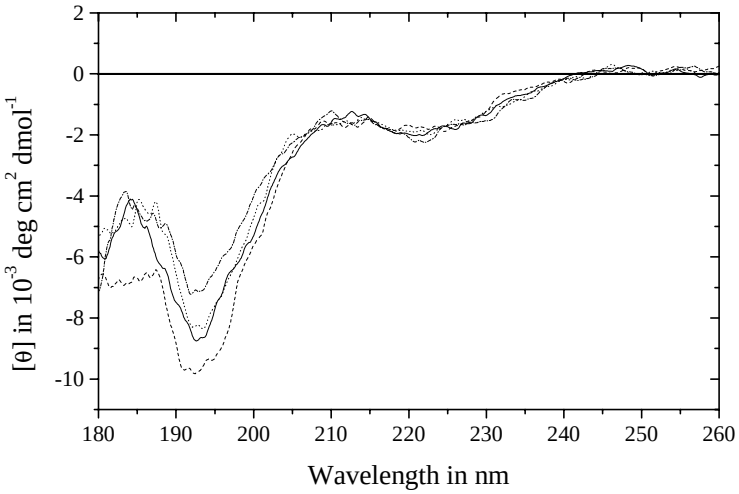


Fig. 35: Temperature effect on the CD spectra of Z-Val-Val-Val-OMe dissolved in HFiP (dashed line: 10 °C, solid: 20 °C, dotted: 30 °C, dot-dashed: 40 °C)

Toniolo et al.[198, 199, 200] have studied a series of Boc- protected oligopeptides Boc-(L-Xaa)$_n$-OMe with n=2-7 as films casts from trifluoroethanol, TFE. Despite the disturbance in the CD spectra due to the overlapping with the intense $\pi \rightarrow \pi^*$ transition of the phenyl ring the similarities between the results of the Italian researchers allow the conclusion on an unordered orientation of the tripeptide derivatives in HFiP solution.

B.3.2.2.5 Peculiarities in the spectra of tripeptide derivatives dissolved in carbon tetrachloride

B.3.2.2.5.1 NH region of selected tripeptide derivatives

Selected cases of dipeptide and tripeptide derivatives are soluble in apolar solvents like n-hexane and carbon tetrachloride. Since the dielectric parameters of these solvents are closer to the properties of a biological membrane the characteristic IR region should be discussed. Especially the tripeptide derivatives containing Ala, Val and Phe amino acids show strong differences in their solubility in apolar solvents. Thus, Z-Ala-Val-Phe-OMe is insoluble in carbon tetrachloride whereas Z-Ala-Phe-Val-OMe and Z-Phe-Ala-Val-OMe are sufficiently soluble up to concentrations of $2*10^{-2}$ mol/l. This would not have been expected from the hydrophobicity parameters discussed in the HPLC chapter of this work. The CHI values measured of the immobilised artificial membrane was 26.07 and 25.87 for Z-Ala-Val-Phe-OMe and Z-Ala-Phe-Val-OMe. Furthermore, the soluble derivatives show a completely

different NH band profile (Fig. 36) which indicates the existence of numerous associates in the case of Z-Ala-Phe-Val-OMe.

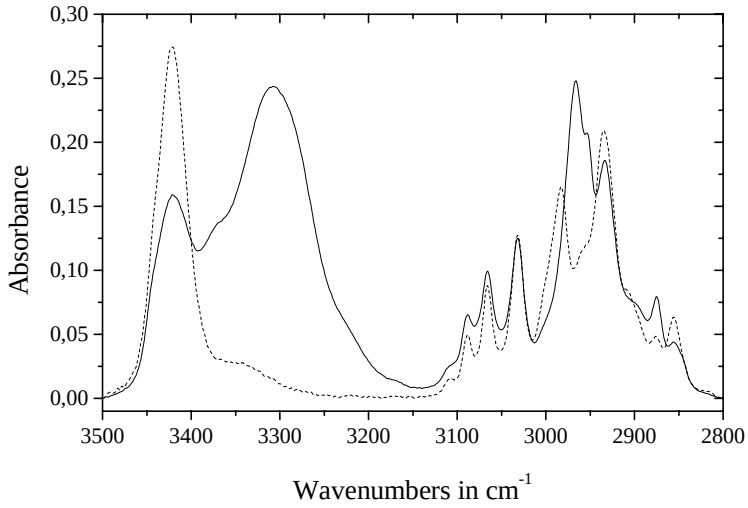


Fig. 36: Spectra of the NH and CH fundamental region of Z-Ala-Phe-Val-OMe (solid) and Z-Phe-Ala-Val-OMe (dashed) dissolved in CCl_4

The NH band profile of the diastereoisomers of Z-Ala-Phe-Val-OMe suggest the occurrence of different associates which result in the complex NH band profile at about 3300 cm^{-1}. Since intermolecular association can be excluded from concentration dependent measurements the associated NH species have to be assigned to intramolecular hydrogen bonds. Those associates are sensitive to the change of the temperature and experience a strong intensity loss on increasing of the temperature. Furthermore, the diastereoisomers show different band profiles.

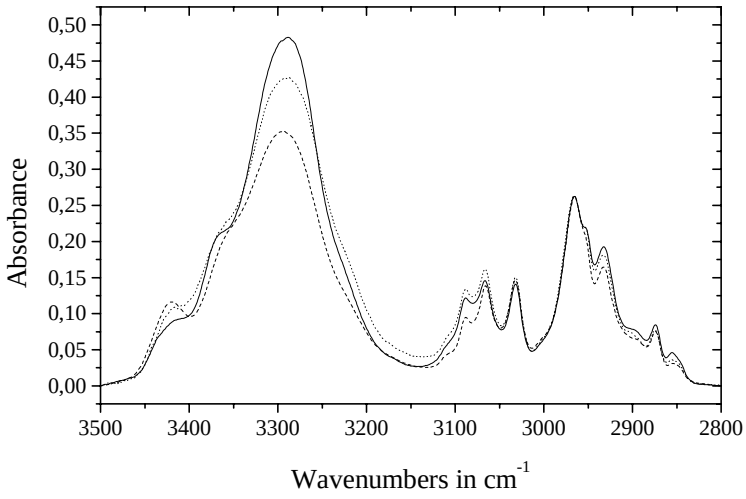


Fig. 37: Spectra of the NH and CH fundamental region of the diastereoisomers of Z-Ala-Phe-Val-OMe (solid: LLL dashed: LDL and dotted: DLL),dissolved in CCl_4 at -10 °C

Band deconvolution of the NH profiles allows the identification of the number of overlapping signals. The temperature dependence enables the assignment of the absorption signals as due to associated and free NH species (Table 36).

Table 36: Absorption maxima of tripeptide derivatives after band deconvolution and peak fitting [in cm^{-1}]

Substance	$\nu_{NH\ free}$	ν_{C5}	$\nu_{ass\ 1}$	$\nu_{ass\ 2}$
Z-Ala-Phe-Val-OMe				
LLL	3444.2	3420.4	3373.9	3311.2
LDL	[a]	3423.7	3365.0	3313.7
DLL	3443.3	3424.6	3388.9	3315.2
Z-Phe-Ala-Val-OMe	3443.6	3421.5	-	3346.5
Z-Ala-Leu-Val-OMe,				
LLL	[a]	3429.9	3380.8	3308.5
LDL	[a]	3426.0	3368.8	3315.7

[a] The intensity of the free signal was too low to locate precisely.

According to the theoretical calculations of Schäfer et al.[201] the size of the associate ring formed by the intramolecular hydrogen bond is a function of the chain length. Hence oligopeptides with up to four amino acid units are inclined to form rings with a maximum number of members, but the stability of a C_{10} ring turns out to be slightly lower than that of a C_7 ring even for tripeptide derivatives. $\Delta\nu$ can be taken as a measure for the energy

stabilisation of the molecule as a result of hydrogen bond formation. In this regard C_7 and C_{10} arrangements will be similar. From the size of the molecule numerous C_7 stabilised conformations are possible and the band profile seen in the NH and CO region will be a product of a distribution of different species.

B.3.2.2.5.2 Amide I region of selected di- and tripeptide derivatives

The carbonyl and amide I region of the tripeptide solutions substantiates the observation of the NH region. The effect of sequence changes results in a rather complex band profile of the diastereoisomers of Z-Ala-Phe-Val-OMe whereas Z-Phe-Ala-Val-OMe shows a broad amide I signal with a band maximum at similar position as found in methylene chloride (Table 37). In methylene chloride band fitting gives band maxima at 1735.6, 1719.0 and 1680.5 cm^{-1} for Z-Phe-Ala-Val-OMe. It should be noted that the analogue derivative Z-Ala-Val-Phe-OMe give maxima at 1743.5, 1718.9, 1688.4 and 1672.8 cm^{-1}, for the ester carbonyl, the urethane and two peptide signals, respectively, and, thus, is rather similar to Z-Ala-Phe-Val-OMe where the corresponding signals were located at 1740.5, 1720.2, 1687.4 and 1670.4 cm^{-1}, respectively.

Table 37: Band positions of the ester and urethane CO band and the amide I signals of di- and tripeptide derivatives dissolved in CCl_4 at 25 °C [in cm^{-1}]

Substances	ν_{Ester}	$\nu_{Urethane}$	$\nu_{Urethane\ ass}$	$\nu_{Peptide\ I}$	$\nu_{Peptide\ II}$	$\nu_{Peptide\ II\ ass}$
Z-Ala-Phe-OMe						
LL	1746.6	1726.4	1711.9	1688.2		-
DL	1746.6	1727.4	1712.0	1687.0		-
Z-Phe-Val-OMe	1744.1	1721.5	-	1687.8	1664.3	
Z-Ala-Phe-Val-OMe						
LLL	1745.1	1725.6	1706.2	1691.2	1674.1	1653.6
LDL	1742.5	1728.0	1707.8	1692.9	1673.4	1656.0
DLL	1745.1	1726.4	1708.0	1691.0	1674.7	1657.5
Z-Phe-Ala-Val-OMe	1740.4	1720.7	-	1684.2		-

The sequence dependence behaviour of the tripeptide derivatives can be traced back to the existence of conformers in solution due to the rotation C^{α}-C' bond axis[166]. Baron et al.[202] have performed investigations on N- and C- protected oligopeptides with two to five amino acid residues containing Val, Ile and Phe. Independently on the amino acid the CO peptide function absorbs at the same frequency. According to the authors the different conformers are

due to the rotation C^{α}-C' bond axis which is hindered if bulky residues are close to each other. Therefore, also in the case of Z-Phe-Val-OMe two amide I signals are observed in solution.

The solvent effect of the conformer equilibrium of Ac-Phe-Val-OMe was studied by Kobayashi[203] and Bystrov[204]. Hence the benzyl function of Phe seems to play an important role in the preference of the rotamers I and II in chloroform (Fig. 38).

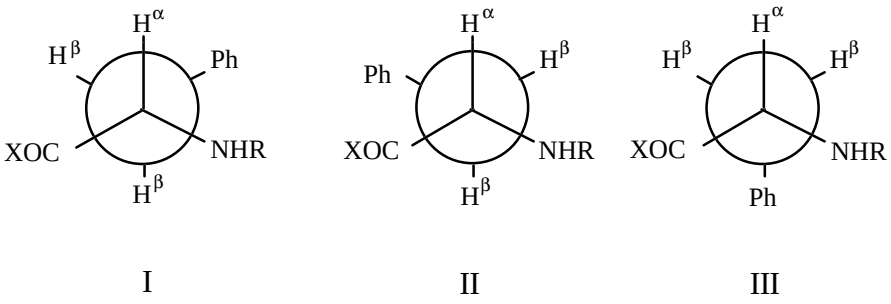


Fig. 38: Sketch of the rotamers resulting form the rotation on the C^{α}-C'-bond axis of Phe

In terms of the predictability of the partitioning behaviour of such compounds from fragmental values this seems an important observation. The sequence of these small oligopeptides determines the solubility in apolar solvents. It is not overestimated to assume that also the partitioning behaviour will depend on the sequence of the peptides. In order to assign the amide I signals of the diastereoisomers of Z-Ala-Phe-Val-OMe VCD measurements were performed.

VCD measurements: The VCD spectra of the amide I region of the diastereoisomers dissolved in carbon tetrachloride is shown in Fig. 39.

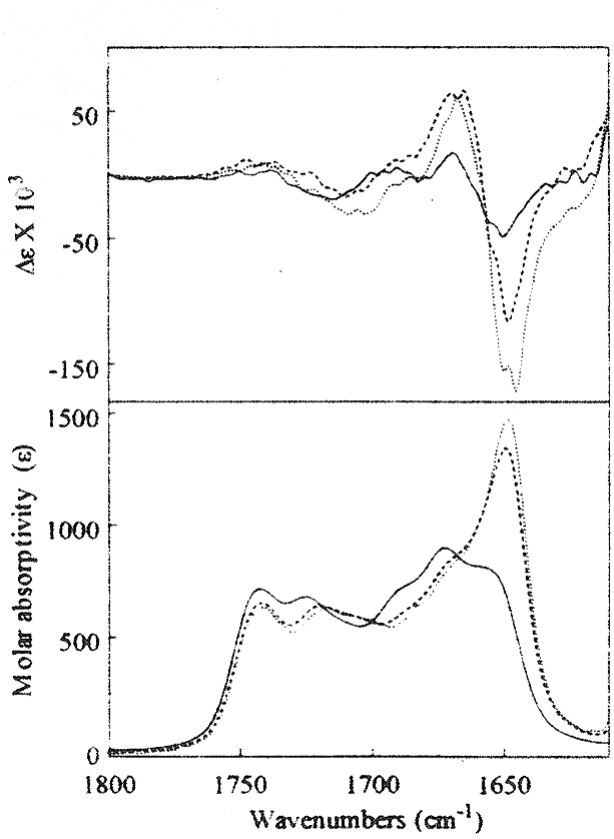


Fig. 39: Overlay of IR and VCD spectra in the amide I region of the diastereoisomers of Z-Ala-Phe-Val-OMe (solid: LDL, dashed: DLL, dotted: LLL) dissolved in CCl_4

To identify the conformers present in solution *ab initio* magnetic field perturbation (MFP) calculations of optimised geometries were carried out on the model compound MeOCO-Ala-Ala-NMe of LL, LD and DL configurations[205]. As discussed in the literature the MFP method give higher agreement between the simulated VCD spectra and the experimental data for defined secondary structure of peptides[206, 207]. Conformations probed included C_5-C_5, C_5-C_7, C_7-C_5, C_7-C_7 and C_{10} ring intramolecularly hydrogen bonded structures. The calculations were performed with *cis* and *trans* arrangement of the urethane function.

Of the eleven stabile LL conformers the three lowest energy conformers were the *trans* C_7-C_7 containing two C_7 rings, (0 kJ/mol), the *trans* C_5-C_7 which combines two intramolecular hydrogen bonded rings of different size within a structure (2.93 kJ/mol) and the *trans* extended β- sheet with two C_5 rings stabilising the extended arrangement (1.67 kJ/mol). The remaining conformers are 7.95 to 50.2 kJ/mol higher in energy. The combined existence of associate rings is not surprising since the interaction on an amide CO function with an NH group in the direct neighbourhood becomes more basic and reacts as an acceptor function with formation of the second hydrogen bond exhibiting a co-operative effect[133]. In dependence on the size of the formed associate ring the entropy loss will make them highly instable with

increasing temperature. Applying this result on Z-Ala-Phe-Val-OMe (LLL) the energetically favoured conformers can be shown in Fig. 40.

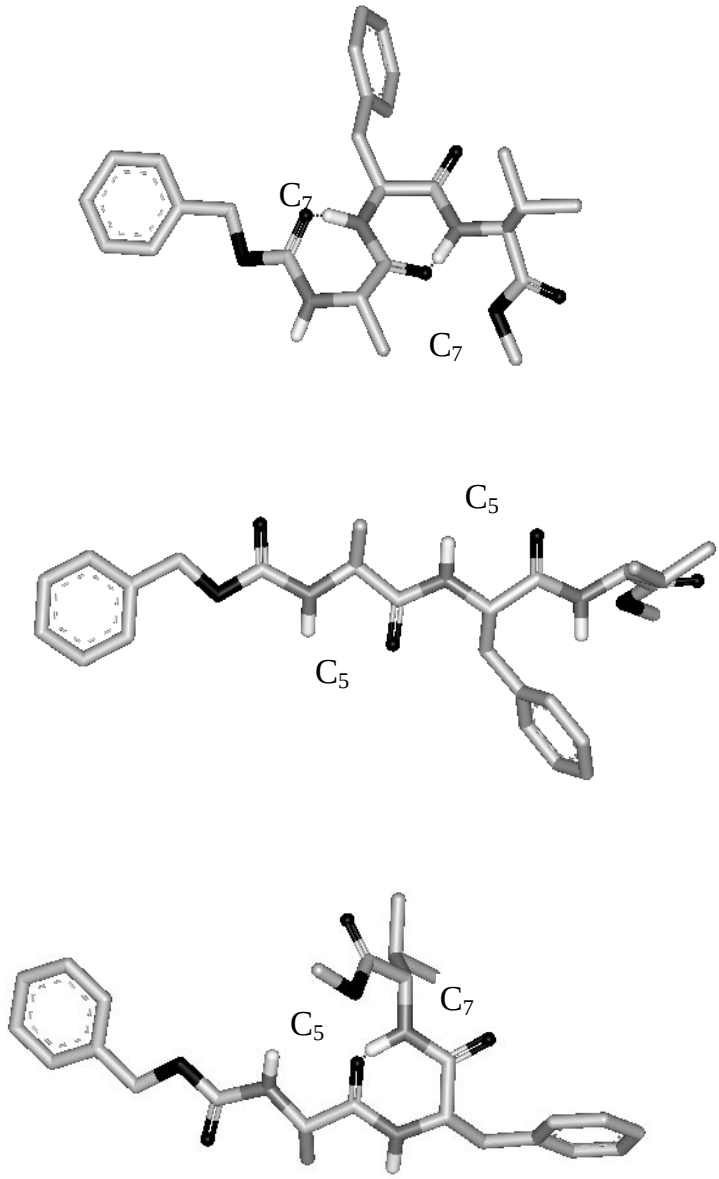


Fig. 40: Low energy conformers of Z-Ala-Phe-Val-OMe (LLL) (above: *trans* C_7-C_7, middle: *trans* extended β- sheet, below: *trans* C_5-C_7) (colour code: black: oxygen, grey: nitrogen, light grey: carbon, white: hydrogen in NH; otherwise, the hydrogen atoms are not shown)

Based on the energy optimised structures the VCD spectra were calculated which correlated well with the experimental data shown in Fig. 39. The *trans* C_7-C_5 rings show a negative VCD band at low frequencies, all three conformers have a positive VCD band at higher frequencies in agreement with the experiment. The calculated IR spectrum gives an intense low frequency band for the *trans* extended β-sheet (Fig. 41).

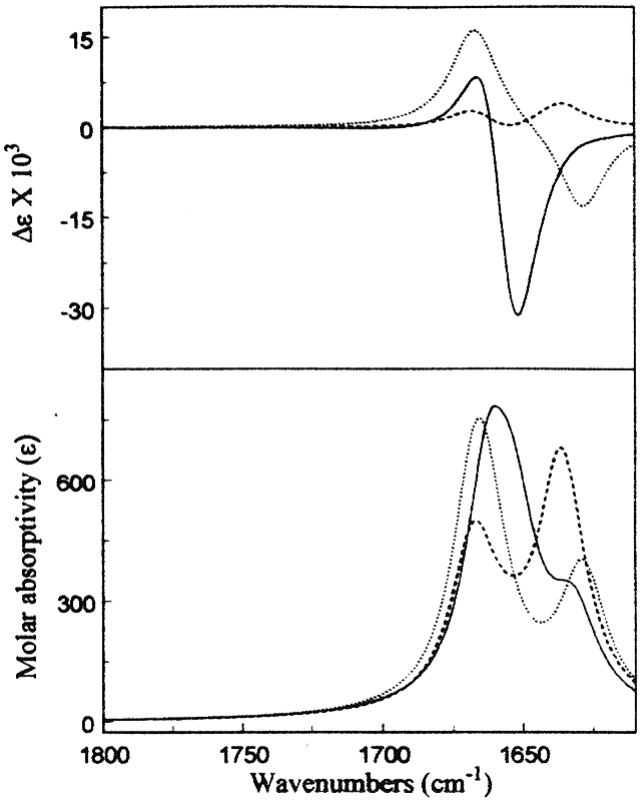


Fig. 41: Calculated IR and VCD spectra of the amide I region of MeOCO-Ala-Ala-NMe. Frequencies are scaled by 0.85 for comparison with the experiment (solid: *trans* C_7-C_7, dashed: extended β- sheet, dotted: *trans* C_5-C_7)

In terms of the diastereoisomers of Z-Ala-Phe-Val-OMe the intense negative VCD band observed at ca. 1650 cm^{-1} which is due to the *trans* C_7-C_7 arrangement causes the main difference in the association behaviour of the compounds. Obviously, the stabilisation of this arrangement is hindered by the chirality of the amino acids. As it would have been expected from the bulkiness of the phenylalanyl residue the tendency to form such associates is lowest for the LDL diastereoisomer since the amino acid side chains of the alanyl and the phenylalanyl residue will be not alternating above and below to the plane of the associate ring but on the same side. Therefore, the repulsion of the residues will lower the population of this arrangement.

Furthermore, for the LD and the DL configuration the occurrence of C_{10} associate rings becomes relevant. Thus, the three lowest energy conformers for the LD configuration of MeOCO-Ala-Ala-NMe were the *trans* C_7-C_7 (0 kJ/mol), *trans* type II β-turn (C_{10}, 2.1 kJ/mol) and the *trans* extended β-sheet (C_5-C_5, 2.93 kJ/mol). For the DL diastereoisomers several conformers have been found within 4.18 kJ/mol of energy: the *trans* C_7-C_7 (0 kJ/mol), the

trans type I β-turn (C_{10}, 2.09 kJ/mol), the *trans* extended β- sheet (C_5-C_5, 2.93 kJ/mol) and the *trans* C_5-C_7 rings (3.35 kJ/mol).

B.3.2.3 Conformational behaviour of peptide derivatives in phospholipids

Immobilised artificial membrane columns (IAM) were introduced into chromatography in order to mimic a cell membrane. In IAM columns the silicate substrate is covered by lecithin molecules which are linked via an amide bond at hydrophobic side of the lecithin functions. In our particular case phosphocholine with two myristoyl esters cover the silicate substrate. As the comparison of the coefficients between the IAM column and biological systems have shown the parameters are not comparable with those of the blood-brain barrier or the skin-water partitioning systems although the surface density of lecithin with 67-77 $Å^2$/molecule is rather similar to the molecular packing in biological membranes[208]. Obviously the differences are caused by the way the lecithin residues are linked to the solid surface: Unlike cell membranes the lecithin molecules of the bonded phase are lacking lateral mobility. Furthermore, the lecithin layer on the solid support is only a monolayer whereas cell membranes consist of a unique mixture of lipids forming a bilayer.

Nevertheless, it will be interesting to observe if the conformation of the tripeptide derivatives is comparable with the conformation in the tested solvents. Furthermore, it could be shown if the presence of the peptide molecules influences the chain mobility of the model membrane. From infrared spectroscopic studies it will be possible to conclude if the peptide derivatives will penetrate in the hydrophobic part of the layer or rather interact with the polar head function.

Phosphocholines have been well characterised by several physicochemical methods in the last decades. One of the most common and hence best investigated saturated diacyl phosphatidylcholine is 1.2-dipalmitoyl-sn-glycero-3-phosphatidyl choline (DPPC)[209, 210] which shows a phase transition from the gel P_β to the lamellar L_α phase at ca. 42 °C. The phase transition is mainly induced by the increasing mobility of the acyl chain (chain melting) which can be monitored on the hand of the red-shift of the antisymmetric and symmetric CH_2 stretching vibration in the infrared spectra indicating a sudden increase of the gauche conformers in the molecular arrangement. Pretransitions occurring at ca. 34 °C (L_β-P_β) were connected to changes in the acyl chain orientation as well but were only observed in a few cases by means of infrared spectroscopy[211]. The low enthalpy values of the so called sub-subtransition (SGII-$L_{\beta'}$) at even lower temperatures (T=7.1 °C) considered as related to a

dehydration of the phospholipid[212] cannot be detected in the IR spectra. Fig. 42 shows the temperature behaviour of the CH_2 symmetric stretching signal with the course of the temperature for selected phosphocholines. As seen for DMPC where the chain melting occurs at ca. 23 °C the shortening of the acyl chain length lowers the main phase transition. Slight differences to the phase transition temperatures may be caused by the use of D_2O for the hydratisation.

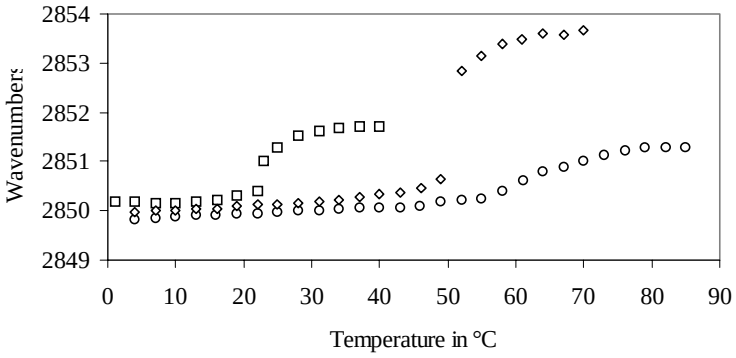


Fig. 42: Change of the CH_2 symmetric stretching vibration of hydrated phospholipids in dependence on the temperature (diamond: DPPC, square: DMPC, circle: DHPC)

DPPC and DHPC show similar thermotropic behaviour with only slight differences in the structure of the gel and subgel phases with interdigitated alkyl chains in 1.2-dihexadecyl phosphatidyl choline (DHPC)[213, 214]. However, linking the alkyl chain to the choline function via an ether bond instead of an ester group causes a slight increase of the lamellar gel-lamellar liquid crystalline phase transition. In our example a smooth phase transition was observed with a similar T_m as observed for DPPC.

Regarding to the influence of amino acids except Trp and Lys on the phase transition temperature of fully hydrated DPPC no significant effect on the chain melting was observed[215]. For our investigations we have chosen DHPC as a lipid model because it allows a simultaneous monitoring of the amide I region and the CH_2 stretching regions in the D_2O hydrated states.

1:1 mixtures of DHPC and the tripeptide derivatives consisting of Ala, Phe and Val amino acid residues were dissolved in chloroform: methanol (v/v 1:1) and spread on an ZnSe-ATR crystal.

Fig. 43 shows the behaviour of the CH_2 symmetric stretching mode with increasing temperature of the hydrated 1:1 mixtures of DHPC and Z-Ala-Phe-Val-OMe, Z-Ala-Val-Phe-OMe and Z-Phe-Ala-Val-OMe.

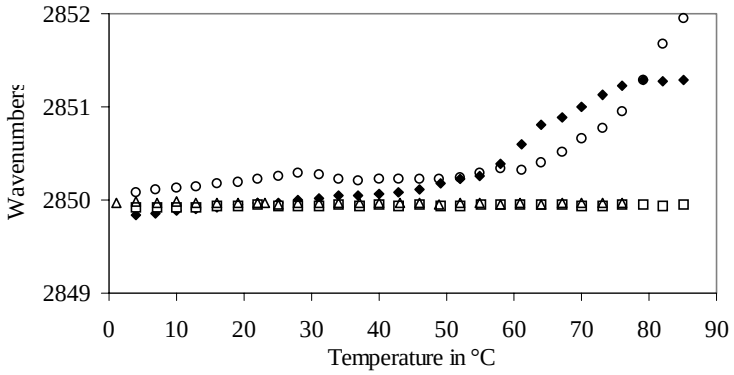


Fig. 43: Change of the CH_2 symmetric stretching vibration of DHPC:tripeptide mixtures vs. the temperature (solid diamond: DHPC, square: DHPC:Z-Ala-Val-Phe-OMe, circle: DHPC:Z-Phe-Ala-Val-OMe, triangle: DHPC:Z-Ala-Phe-Val-OMe)

As clearly seen in the picture the presence of a tripeptide has major consequences on the phase transition behaviour of DHPC. Furthermore, the effects are dependent on the sequence of the tripeptide derivative. Whereas the presence of Z-Ala-Phe-Val-OMe or Z-Ala-Val-Phe-OMe results in the disappearance of a lamellar L_α phase in DHPC the interaction of the phospholipid with Z-Phe-Ala-Val-OMe induces a different structure which is characterised by a higher amount of *gauche* conformers and a higher T_m than in the pure DHPC. From the IR spectra the chain melting temperature can be approximated at ca. 65 °C in the system containing Z-Phe-Ala-Val-OMe.

From the amide I region the properties of the tripeptide derivatives can be concluded (Table 37).

Table 37: Positions of the ester carbonyl, urethane and peptide signals of the hydrated DHPC: tripeptide mixtures at 22 °C [in cm^{-1}]

Substance	ν_{Ester}	$\nu_{Urethane}$	$\nu_{Peptide}$
DHPC: Z-Ala-Phe-Val-OMe	1741.7	1680.8	1636.6
KBr:	1742.2	1687.0	1644.0
DHPC: Z-Ala-Val-Phe-OMe			
KBr:	1747.8	1686.2	1643.9
	1739.8		1626.7
DHPC: Z-Phe-Ala-Val-OMe	1748.8	1702.8	1675.3
			1656.0
KBr	1754.0	1696.2	1675.3
	1747.7		1658.7

For comparison the positions of the corresponding bands in the KBr spectra are given. All tripeptide derivatives are insoluble in water. However, the band profiles are not similar to those found in apolar solutions either. Conclusively in the presence of water the tripeptide molecules will occur in a crystalline like orientation which does however strongly influence the phase behaviour of the phospholipid.

In the absence of water the amide I region of DHPC:Z-Ala-Phe-Val-OMe shows a different band pattern (Fig. 44) which is rather similar to Fig. 38 showing Z-Ala-Phe-Val-OMe dissolved in carbon tetrachloride whereas the amide I band profile of the Z-Phe-Ala-Val-OMe and Z-Ala-Val-Phe-OMe mixtures exhibit the features of the crystalline state.

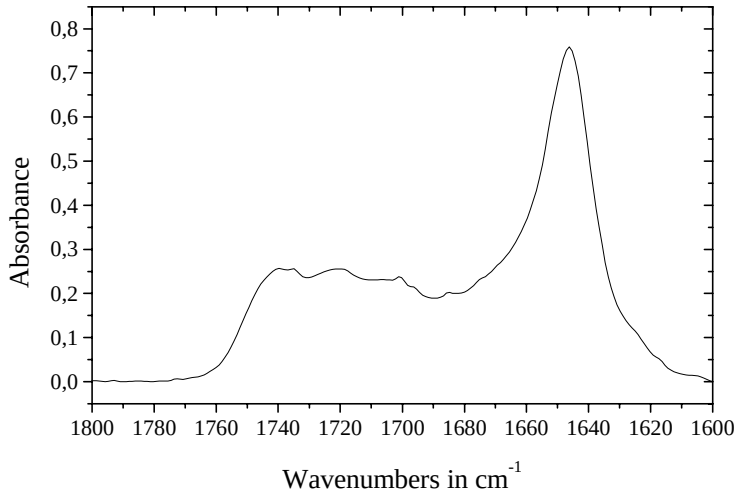


Fig. 44: Amide I region of the DHPC:Z-Ala-Phe-Val-OMe mixtures (dry state) at 25 °C

Hence only Z-Ala-Phe-Val-OMe will be enclosed into the hydrophobic part of the phosphocholine. For the other systems phase separation is more likely.

Another peculiarity of the system containing Z-Phe-Ala-Val-OMe should be mentioned here. As shown in the Fig. 45 the band profiles in the region between 1800 and 1600 cm^{-1} changes in the course of the temperature which indicates changes in the crystalline structure of the tripeptide derivative. The melting temperature of the pure substance is found at 125 °C. This reversible effect may induce the effects discussed in Fig. 43.

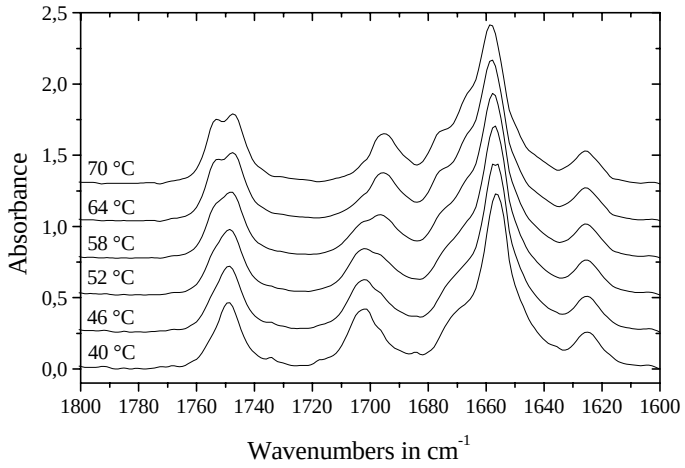


Fig. 45: IR spectra in the region between 1800 and 1600 cm^{-1} of DHPC:Z-Phe-Ala-Val-OMe at different temperatures

Contrary to the observations of Jacobs and White[216, 217] in DMPC:Ala-Xaa-Ala-OtBu (Xaa= Gly, Ala, Phe, Trp) systems the tripeptide derivatives do not simply act as impurities but induce a strong perturbation in the DHPC layers.

B.3.3 Intramolecular hydrogen bonding in ureido sugar amino acid and dipeptide derivatives

As we have shown before the title compounds which are intermediates in the synthesis of effective anticancer drugs can be characterised by QSPR descriptors from chromatographic gradient measurements. Furthermore it was shown that these values can be predicted on the hand of fragmental amino acid descriptors with satisfying accuracy. In some cases the experimental effective overall hydrogen bond basicity $\sum\beta_2^H$ received from amino acid fragmental values did not give satisfying results when compared with the predicted descriptors based on amino acid fragmental descriptors. In particular this was found for the derivatives with Gly-OEt, D-Val-OEt, Phe-Gly-OEt and Ala-Phe-OBzl residues in the side chain. In a former study we have observed that some ureido-2-deoxy-β-D-glucopyranosides behave exceptional regarding the H/D exchange rate on N-1-H and N-3-H in the NMR experiment[218]. At a first idea it was believed that both NH function would differ in their hydrogen bond basicity and give the ^1H chemical shift values in table 38. However, the comparison of the net charges on both NH groups of the valyl and alanyl derivatives of the ureido sugars refute this assumption.

Table 38: Net charges at the N and H atoms as found by AM1 calculations and the ^{1}H chemical shift of the NH functions

	Charges on		Charges on		^{1}H chemical shift[219] at	
	N-1	H at N-1-H	N-3	H at N-3-H	N-1-H	N-3-H
Ala	-0.3717	0.2430	-0.3594	0.2270	4.47	5.29
Val	-0.3834	0.2416	-0.3908	0.2303	4.53	5.70

Hence conformational peculiarities in these systems need to be considered because it is known that the glycoside residues affect the peptide conformation. In the literature two opposite mechanism have been described for glycopeptides[220]. One approach can be defined as a simple reduction of the variety of conformers of otherwise flexible peptide chains by bending of the peptide backbone away from the glycoside[221]. As a consequence turns are introduced in the peptide residue. The nature of the glycoside can influence the character of the turn[222, 223]. Furthermore, specific interactions e.g. hydrogen bonds between the sugar and the peptide function can result in „glyco-turns“[224]. However, there are also examples in the literature where no changes in the peptide backbone are observed due to glycosylation compared to the parent molecule[225].

Our title compounds are slightly different from the glycopeptides described in the literature since the glycoside is linked to the peptide residue via an ureido function which will enforce a planar arrangement in this part of the molecule. Hence the question arises if the intramolecular hydrogen bonding will still occur and result in a back folding of the peptide backbone.

Indeed, the infrared spectra of the chloroform solution show a complex band profile in the NH stretching region which obviously varies with the amino acid residue and chirality of the amino acid (Fig. 45).

Band deconvolution of the NH profiles gives three NH bands at ca. 3453, 3422 and 3375 cm^{-1} which can be assigned to a free and two intramolecularly associated NH species. $\Delta\nu$ between the free and the low wavenumber signal is equal to 78 cm^{-1}. For Z-Ala-Phe-Val-OMe dissolved in carbon tetrachloride the NH signal of a C_7 associate was shifted by ca. 72 cm^{-1} to lower wavenumbers. Hence the intramolecular hydrogen bond is of similar strength in the ureido sugar and can be assigned to an associate forming a seven membered ring (C_7). Analogously, the intense signal at 3422 cm^{-1} represents a C_5 NH associate.

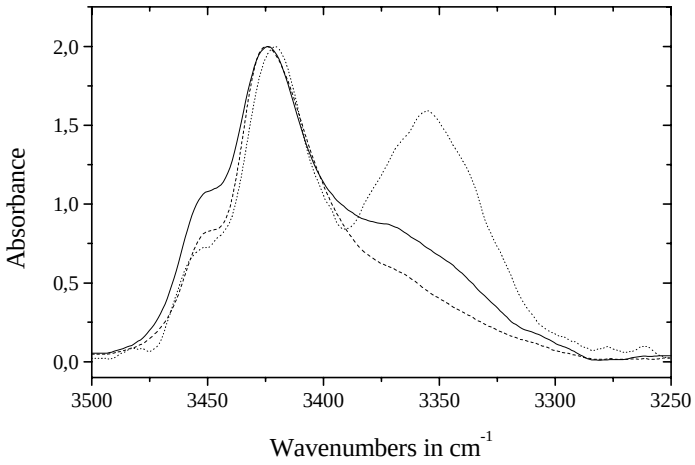


Fig. 46: Spectra of the NH band profile of the ureido sugar derivatives dissolved in chloroform with Gly (solid line), Phe (dashed) and D-Val (dotted) in the side chain

Table 39: Positions [in cm^{-1}] and relative intensities [in %] of free, C_5 and C_7 NH functions at 20 °C

				Relative intensities		
Residue	ν_{NHfree}	ν_{NHC5}	ν_{NHC7}	NH_{free}	NH_{C5}	NH_{C7}
-Gly-OEt	3454	3424	3372	35.4	58.4	6.1
-Ala-OEt	3454	3423	3391	19.8	75.1	5.1
-Val-OEt	3455	3420	3355	22.7	64.7	12.6
-D-Val-OEt	3454	3421	3373	51.6	42.0	6.4
-Leu-OEt	3454	3423	3393	18.8	75.8	6.5
-Phe-OEt	3453	3423	3395	25.0	68.5	6.5
-Gly-Ala-OEt	3453	3420	3384	43.6	48.5	7.6
-Gly-Val-OEt [a]	3451	3421	3390	40.5	51.3	8.2
-Gly-Phe-OBzl	3452	3418	3387	29.8	63.1	7.0
-Ala-Gly-OEt	3453	3423	3374	37.3	56.5	6.2
-Val-Gly-OEt [a]	3453	3424	3380	42.3	52.1	5.6
-Leu-Gly-OEt	3453	3424	3382	40.6	53.0	6.3
-Phe-Gly-OEt	3452	3421	3371	39.5	54.9	5.5
-Ala-Ala-OEt	3454	3421	3376	18.6	71.5	9.9
-Ala-Phe-OBzl	3455	3420	3379	15.9	75.6	8.4

[a]Here the identification of the NH signals was ambiguous; also two different C_7 signals would fit the band profile. For the sake of equivalent treatment the C_7 band was represented by only one Voigt function.

In general, the band positions of the dipeptide ureido sugar derivatives are similar to those of the amino acid derivatives allowing small deviations of the band positions due to the mass effect. Differences occur only in the intensities of the separated signals (Table 39).

Although the type of the hydrogen bonds are the same as for the amino acid derivatives the variety of potential hydrogen bond forming functions is increased by the enlargement of the peptide residue. Whereas C_7 associates in the amino acid derivatives can be formed by interaction of the N-1-H and the acetyl function at C-3 of the glucopyranoside an additional C_7 associate would be possible in the dipeptide derivatives between the N-6-H and the ureido carbonyl function (Fig. 47).

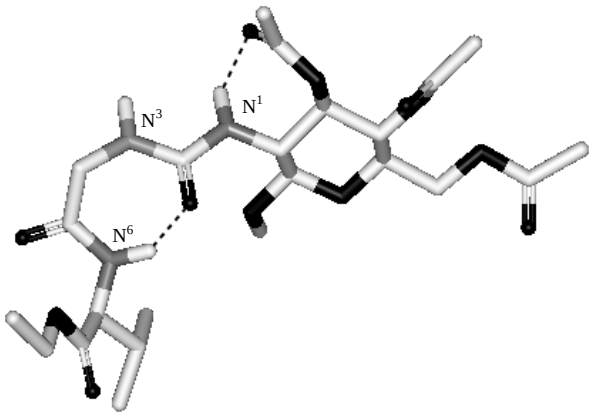


Fig. 47: One possible conformation of the -Gly-Val-OEt ureido sugar derivative. Intramolecular C_7 hydrogen bonds are indicated. (colour code: black: oxygen, grey: nitrogen, light grey: carbon white: hydrogen in NH; otherwise, the hydrogen atoms are not shown)

As the high intensity of the NH signal at 3420 cm^{-1} indicates most species will occur in the extended peptide conformation which is stabilised by C_5 associates involving the N-3-H and N-6-H. Additionally, an interaction of the methoxy function at C-1 and the N-3-H group is theoretically possible but, up to our knowledge, this associate was only observed in the solid phase of ureido sugars with D-valine substitution[226].

In order to judge the concentration of the associated species the extinction coefficient would be needed. However, these values are not available from these spectra. In the tripeptide derivatives the NH extinction coefficient of a C_5 associated group was estimated as ca. 1.5 to 3 times larger than ε_{NHfree}[166]. For NH signals involved in a C_7 associate ring ε_{NHC7} will be even larger and can be approximated as up to 10 times of ε_{NHfree}. Based on these simplifications the relative concentration of non associated, C_5 and C_7 associated NH species are given in table 39 using a factor of 2 and 7 for the correction of the NH extinction coefficients of C_5 and C_7 signals, respectively.

The comparison of the concentration of the NH species in the amino acid derivatives there are two compounds exhibiting a significantly higher amount of free NH groups than found for the other substances, namely for -Gly-OEt and -D-Val-OEt. As cited above the steric aspects of the D-valyl residue might hinder a hydrogen bond to the acetyl function at C-3 but may only be able to stabilise weaker associates with the methoxy function on C-1.

Due to its structural peculiarity the amino acid Gly is not restricted in its mobility. Hence the peptide backbone is more flexible and not influenced by repulsion of the side chain and therefore, not even supports the formation of C_5 hydrogen bonds.

Outstanding with regards to the tendency to form C_7 associates are the data of -L-Val-OEt. Obviously the branched side chain of the valyl residue supports the intramolecular hydrogen bond with the acetyl function at C-3. This has significant consequences on the accessibility of the N-3-H in the H/D exchange NMR experiment[227] but has obviously no significant effect on the solute descriptors.

The data of the dipeptide ureido sugar derivatives in the last two columns show that the percentage of free and C_5 associated NH species is of the same order for -Xaa-Gly and -Gly-Xaa sequences. Thus, the peptide chain is mainly in the extended conformation which is even more favoured in the cases of the -Ala-Ala-OEt and the -Ala-Phe-OBzl derivatives. Obviously, the molecules are forced to arrange in the extended form due to repulsion of the bulky residues on the C-α, although the C_7 associates would be enthalpically more stable. The low amount of free NH functions might lead to the large positive deviation in the correlation of β_2^H for -Ala-Phe-OBzl derivatives.

Table 40: Temperature coefficients $\Delta\nu/\Delta T$ of the NH signals [in 10^{-3} cm^{-1}/K].

Residues	$\Delta\nu/\Delta T$ of NH ($_{free}$)	$\Delta\nu/\Delta T$ of NH (C_5)	$\Delta\nu/\Delta T$ of NH (C_7)
-Gly-OEt	21.7	41.7	312
-L-Ala-OEt	[a]	1.0	107
-L-Val-OEt	31.4	60.4	478
-D-Val-OEt	7.2	13.0	7.9
-Ala-Gly-OEt	29.3	57.1	299
-Gly-Ala-OEt	11.6	25.7	156
-Val-Gly-OEt	31.4	15.8	153
-Gly-Val-OEt	26.9	13.3	148
-Phe-Gly-OEt	19.0	32.9	170
-Gly-Phe-OEt	[a]	38.0	86

[a]Intensity of the band is very low at -30 °C

The temperature effect on the positions of the NH signals gives an indication about the enthalpy of the associate formation. As expected the shift of the NH position with the temperature is generally very small for the C_5 and the free NH signals but by an order of magnitude larger for the C_7 associates (Table 40).

For the amino acid derivatives we have found a strong dependence $\Delta\nu/\Delta T$ on the kind of amino acid residue. In the case of the dipeptide derivatives the stability of the intramolecular hydrogen bonds seem to follow different rules. Here the amino acid sequence determines $\Delta\nu/\Delta T$. Thus, compounds with -Xaa-Gly sequences (Xaa = Ala, Val, Phe) have a stronger temperature effect than the corresponding -Gly-Xaa derivatives. Obviously, the bulkiness of the amino acid Xaa lowers the enthalpy effect of stable C_7 associates. -Gly-Xaa sequences are less affected; the bulky residue of the second amino acid might repel with the acetyl functions at the glycoside in the backfolded conformer as sketched in Fig. 46. Thus, the C_7 type of intramolecular hydrogen bonds are not supported in -Gly-Xaa sequences and the dipeptidyl function favours the extended conformation. When discussing the data of the ureido sugar with the amino acid combination of Gly and Val it seems surprising that $\Delta\nu/\Delta T$ for the free NH is larger than for the C_5 signal. However, we have to recall that the signal defined as free NH is indeed a result of three different NH functions. What is observed here might be also a result of the different temperature behaviour of the associates because the acidity of the three NH functions will be slightly different although the NH signal can not be resolved sufficiently by the IR technique.

The strength of the formed hydrogen bond and its stability against temperature variation has not only an effect on the deshilding properties of the NH signals[228, 229, 230] which will affect the H/D exchange rate proved by ^{1}H NMR experiments[231] but is also relevant for synthetic purposes which include the partial introduction of a nitroso function on N-1-H. Furthermore, the differences in the intramolecular hydrogen bonding behaviour seem to reason the deviation between predicted and experimentally determined descriptors in N-(2-amino-2-deoxy-β-D-glucopyranoside)-N'-carbamoyl-L-amino- acid- and -dipeptidyl esters.

B.3.4 Quantitative analysis of the association behaviour

The calculation of equilibrium constants for the interaction of a proton donor A and a proton acceptor B from infrared measurements bases on the observation of the decrease of the

intensity of a sensitive signal like the OH or NH stretching vibration which is caused by the formation of associates.

$$K = \frac{c_{ass}}{(c_{B0} - c_{ass}) \cdot c_A} \qquad [16]$$

c_{ass}= $c_{A0} - c_A$
c_{A0} = initial concentration of the donor compound
c_{B0} = initial concentration of the acceptor compound

Using infrared spectroscopy for the determination of an equilibrium constant K for the association of an alcohol the Lambert-Beer`s law can be applied and K_{OH} is[232]

$$K_{OH} = \frac{E_{A0} - E_A}{E_A \cdot \left(c_{B0} - c_{A0} \cdot \left[\frac{E_{A0} - E_A}{E_{A0}} \right] \right)} \qquad [17]$$

E_{A0} absorbance before the addition of the acceptor compound
E_A absorbance after the addition of the acceptor compound
c_{B0} initial concentration of the acceptor compound
c_{A0} initial concentration of the donor compound

However, this equation is only valid when 1:1 association occurs and no self-association of the alcohol molecules can be observed. Furthermore, the interacting donor molecules should only exhibit one active acceptor (donor) function. This condition limits the reliability of eq. 16. Among others Zeegers-Huyskens et al.[233, 234] attempted to recognise the peculiarities of the hydrogen bond formation of polyfunctional hetereocyclic bases with phenol and set up eq. 18 which gives results with a standard deviation of 10%.

$$\frac{1}{2 \cdot c_A + K_{app} \cdot c_A^2} = \frac{1}{K} \cdot \frac{K_{app}}{2 \cdot c_A + K_{app} \cdot c_A^2} - K_{1:2} \qquad [18]$$

c_A	concentration of the monofunctional proton donor
K_{app}	apparent equilibrium constant, calculated according to eq. 17
$K_{1:2}$	equilibrium constant for the formation of 1:2 associates

This equation works very well for molecules where the acceptor functions are similar in their hydrogen bond basicity and no intramolecular hydrogen bond formation is possible. Furthermore, it is essential that the acceptor functions in the molecules are accessible without difficulties and the 1:1 associated species do not influence the second association. The conditions for this approach are not fulfilled for our peptide molecules. In contrast to the flexible peptide backbone different associates are possible which might influence each other.

In order to gain more information about associate formation which cannot be treated as 1:1 associates we have studied a selection of diols as model systems for polyfunctional proton donors. Although the peptides will mainly act as proton acceptors similar rules should be obeyed.

B.3.4.1 Hydrogen bond formation in alcohols and diols dissolved in methylene chloride

A selection of diols were analysed regarding their capability to form intermolecular hydrogen bonds to heterocyclic proton acceptors such as pyridine, 2.4-lutidine and 2.4.6-collidine. According to the UCL database[74] the effective hydrogen bond basicity of these acceptor molecules is 0.52, 0.63 and 0.70, respectively.

The solute descriptors of propyl alcohols taken from the UCL database are given table 41.

Table 41: Solute descriptors of propanols, propandiols and glycerol [V_x in 10^{-2} l/mol]

Substance	R_2	π_2^H	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x	$\log P_{Oct}$
Propan-1-ol	0.236	0.42	0.37	0.48	0.5900	0.25
Propan-2-ol	0.212	0.36	0.33	0.56	0.5900	0.05
Propan-1.2-diol	0.373	0.90	0.58	0.80	0.6487	-0.92
Propan-1.3-diol	0.397	0.91	0.77	0.85	0.6487	-1.04
Glycerol	0.512	0.9	0.7	1.14	0.7074	-1.76

Regarding to the hydrogen bond properties of these molecules it can be recognised that the hydrogen bond acidity is lower in the secondary than in the primary alcohol. Obviously, the steric hindrance for a ideal hydrogen bond geometry become noticeable. Regarding their proton donor capabilities propan-1.2-diol and propan-1.3-diol could be treated as the sum of the effective hydrogen bond acidity of propan-1-ol and propan-2-ol and the sum of two propan-1-ol residues. The simple arithmetic gives values of 0.7 and 0.74, for propan-1.2-diol and propan-1.3-diol, respectively. The experiment give values for propan-1.3-diol which are slightly higher and in the case of propan-1.2-diol the effective hydrogen bond acidity is significantly lower than approximated. For glycerol the hydrogen bond acidity is even lower than for the diols expressing clearly that an increase of the number of proton donor function does not necessarily give stronger hydrogen bond donors. Similar conclusions can be drawn when discussing the hydrogen bond basicities. The explanation of the strong deviation from additivity can be found in the formation of intramolecular hydrogen bonds.

B.3.4.1.1 Intramolecular association in diols

Quantum mechanical calculations on propan-1.2-diol and propan-1.3-diol have shown that the intramolecular hydrogen bonds deliver a energetic stabilisation of the molecules which confirmed the results of high resolution microwave studies[235]. Thus, *ab initio* calculations give a energy differences E_r between the *all-trans* arrangement and the most stable conformer of 11 and 28 kJ/mol, for propan-1.2-diol and propan-1.3-diol, respectively[236, 237] Semi-empirical PM3 method leads to similar results even though the stabilisation energy of both alcohols is not as different: -14 and -17 kJ/mol, for propan-1.2-diol and propan-1.3-diol, respectively[238]. The differences in the stabilisation energy are without any doubt caused by the geometry of the intramolecular arrangements which are sketched in Fig. 48 and 49.

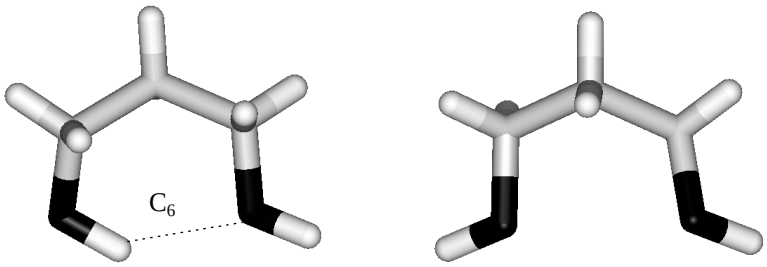


Fig. 48: Conformers of propan-1.3-diol (left: tGG'g conformer, right: tTT't conformer; E_r= 19.09 kJ/mol (colour code: black: oxygen, grey: carbon, white: hydrogen)

The hydrogen bond associate in propan-1.3-diol is characterised by a hydrogen bond distance of 2.02 Å and an angle β of 47°[239].

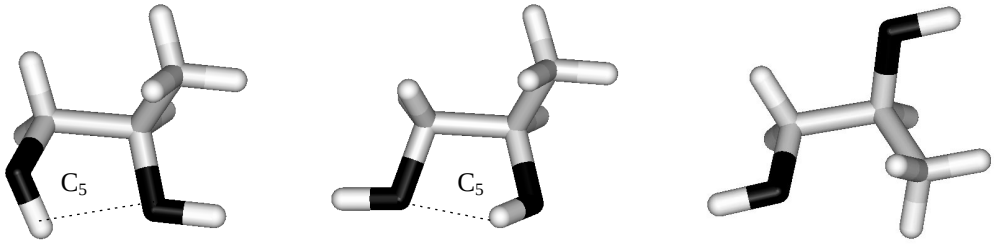


Fig. 49: Conformers o propan-1.2-diol (left: g'Gt; middle: tGg' and right: tTt conformer[240] (colour code: black: oxygen, grey: carbon, white: hydrogen)

In Fig. 49 selected conformers of propan-1.2-diol are shown in order to demonstrate which geometries are found for this diol among numerous others. The relative energies of the three conformers were calculated to 0, 2.5 and 10.9 kJ/mol, for g'Gt, tGg' and tTt', respectively. Although the hydrogen bonds form a five membered associate rings in both arrangements, g'Gt and tGg', due to its hydrogen bond strength the primary OH will preferably act as the proton donor. Contrary the hydrogen bond basicity of a secondary OH function is larger than for a primary OH and makes them a better proton acceptor function. The geometries for the hydrogen bond distance and angle β are similar in both structures and equal to 2.19 Å and 71°.

The stabilisation of conformers by intramolecular hydrogen bonds is dependent on the distances of the interacting functions but also on the hydrogen bond angle β which should become close to zero in the perfect arrangement. Hence the six membered ring gives larger E_r values. However, discussing the problem in terms of entropy the six membered ring is more flexible.

The theoretical consideration will reflect on the spectroscopic behaviour of the molecules. As can be seen in Fig. 50 the OH band region of the diols exhibits an additional broad signal due to the intramolecular hydrogen bonds. Furthermore, the differences of the hydrogen bond acidity discussed above affect the position of the OH signal of the propanols.

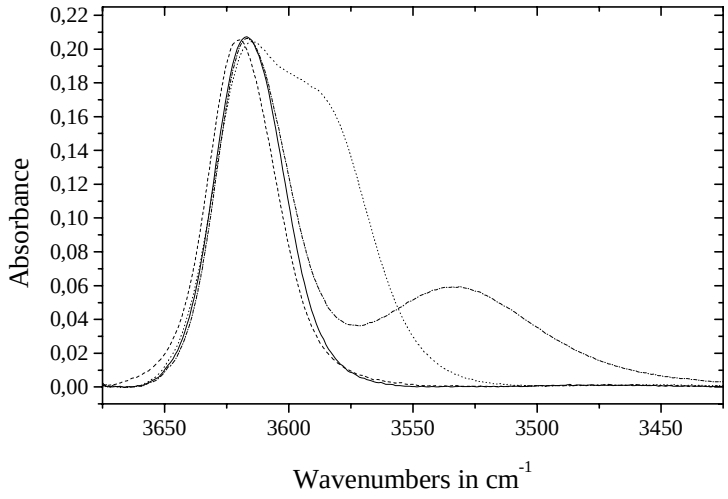


Fig. 50: Normalised OH region of propan-1-ol (solid), propan-2-ol (dash), propan-1.2-diol (dot) and propan-1.3-diol (dash-dotted line) dissolved in methylene chloride at 25 °C

Table 42 gives the OH band positions of the propyl alcohols, their molar extinction coefficients and the temperature dependence of ν and ε_{Mol} based on the assumption of linearity in the temperature range between -20 and 30 °C.

Table 42: OH band positions and the molar extinction coefficients at 25 °C and their temperature dependence as observed after band deconvolution and peak fitting

Substance	ν_{OH} [in cm^{-1}]	$\Delta\nu/\Delta T$ [in cm^{-1}/K]	ε_{Mol} [in l/mol mm]	$\Delta\varepsilon_{Mol}/\Delta T$ [in l/mol mm K]
Propan-1-ol	3616.9	0.1297	6.562	-0.0317
Propan-2-ol	3619.8	0.1279	7.252	-0.0473
Propan-1.2-diol	3619.0	0.090		
	3587.2	0.077		
Propan-1.3-diol	3621.5	0.558		
	3533.7	1.176		

Supported by the relative wavenumber shifts received from *ab initio* calculations[239] the OH signals at 3621.5 and 3533.7 cm^{-1} can be assigned to the free and the intramolecular hydrogen bonded OH species in propan-1.3-diol, respectively. The free OH signals which would belong to the conformers similar to those depict in Fig. 48 cannot be distinguished.

Intramolecular hydrogen bonds are also observed in *ortho* substituted cyclohexandiols. Steric hindrance of the rotational freedom enables the distinction in *trans*- and *cis*-cyclohexan-1.2-diols which also give a different OH band profile in the infrared spectra. In the assignment

of the signals the conformational variety of the cyclohexane ring which would be relevant for cyclohexan-1.4-diols can be neglected[241]. In cases of cyclic 1.2-diols it could be shown that the OH groups were closer and interacting to a larger extent in the *cis* form of cyclohexan-1.2-diol then in the corresponding *trans* form[242]. Hence, based on steric considerations it can be explained why $\Delta\nu_{OH}$ is 33.7 and 29.2 cm^{-1} for the *cis* and the *trans* form, respectively. Due to the cyclic arrangement the axial and equatorial positioning of the OH groups would cause different conformers. However, for *trans*-cyclohexan-1.2-diol there was no evidence for the diaxial form but the diequatorial form would be expected to be stabilised by the intramolecular hydrogen bond leading to the signals given in table 43.

Cis-cyclohexan-1.2-diol may exist in intramolecular forms in which the axial OH may be bonded to the equatorial O or the equatorial OH may be bonded to the axial O. But, even in the matrix isolation spectra a differentiation of both types was not possible.

Table 43: OH band positions and the molar extinction coefficients at 25 °C and their temperature dependence as observed after band deconvolution and peak fitting of cylcohexandiols

Substance	ν_{OH} [in cm^{-1}]	$\Delta\nu/\Delta T$ [in cm^{-1}/K]	ε_{Mol} [in l/mol mm]	$\Delta\varepsilon/\Delta T$ [in l/mol mm K]
Cyclohexan-1.2-diol (*cis*)	3610.9	0.083	7.856	0.0285
	3577.2	0.1939	7.445	0.0122
Cyclohexan-1.2-diol (*trans*)	3612.9	0.0490	6.95	-0.356
	3583.7	0.0695	10.62	-0.016
Cyclohexane-1.4-diol (*rac*.)	3605.1	0.808	16.3	-0.0272

As it was discussed above the hydrogen bond acidity of propan-1.3-diol can be described roughly by acidity of two propan-1-ol units. Hence the estimation of thermodynamic parameters should be possible with the following approximation:

- The acidity of one free OH of the diols is similar to propan-1-ol
- Based on the first approximation the molar absolute extinction coefficients are equal for one free OH signal of the propandiols and the OH band of propan-1-ol.
- The temperature dependence of ε_{OH} of the free OH signals of the diols is equal to $\Delta\varepsilon_{OH}/\Delta T$ of propan-1-ol.
- For the cyclohexan-1.2-diols the correlation was done based on the molar extinction coefficient of cyclohexan-1.4-diol abstracting from the influence of equatorial and axial orientation and *cis/trans* compounds.
- Based on this approximation the concentration of the intramolecular associated OH can be calculated according to Lambert-Beer's law.

The equilibrium constants K for the intramolecular association is given in table 44.

Table 44: Equilibrium constants K at 25 °C for the intramolecular association

Substance	K
Propan-1.2-diol	0.81
Propan-1.3-diol	1.52
Cyclohexan-1.2-diol (*cis*)	1.16
Cyclohexan-1.2-diol (*trans*)	1.31

As expected the results in table 44 indicate that the compounds generally favour the formation of intramolecular associated species over the non-associated structure. Only the equilibrium constant for propan-1.2-diol was smaller than 1.

We have also investigated the hydrogen bonding behaviour of diols in which both OH functions are separated by longer alkyl chains namely butan-1.4-diol, pentan-1.5-diol and hexan-1.6-diol. Among these diols only butan-1.4-diol forms a seven membered intramolecular associate ring which can be identified on the hand of the additional OH stretching signal at 3430 cm^{-1}. Pentan-1.5-diol and hexan-1.6-diol give a symmetric OH stretching signal in the spectra. Hence no intramolecular associates need to be considered. Instead both OH group will associate independently with two times of the hydrogen bond acidity as the corresponding mono alcohols ($\sum\alpha_2^H$: n-hexanol: 0.37; hexan-1.6-diol: 0.75).

B.3.4.1.2 Intermolecular association with pyridine derivatives

To gain results of the thermodynamic parameters of the different association species, namely 1:1 and 1:2 (alcohol :amine) associates we have to introduce an approximation:

At low amine concentrations only 1:1 associations should be relevant for the deductions. From the statistical point of view in the case of 10% associated OH groups only 1% will be associated as 1:2 associate. Based on this approximation the decrease of the free OH band after addition of the amine, ΔE_{OH}, represents the molar quantity of the formed associates, but the actual intensity of the free OH band contains also the free OH groups of the diol species associated on one OH function which is just the concentration represented by ΔE_{OH}. This has to be recognised when calculating the equilibrium constant. The equilibrium constants of association process may be extrapolated to zero concentration of amine, but in most cases the constants do not change, if the concentration of associates is lower than 10% of the alcohol molecules. This constant is defined as $K_{1:1}$.

At high amine concentrations the equilibrium between free molecules and 1:1 associates keeps constant but now 1:2 associates will be formed and need to be considered in the calculation of equilibrium constants. Now the free OH band represents the concentration of the non-associated alcohol molecules and the free OH groups of the 1:1 associates. The concentration of the 1:1 associates is given via the correlation with $K_{1:1}$, which allows to determine $\Delta E_{OH\ 1:1}$. Conclusively the actual loss of intensity of the free OH band ΔE_{OH} is equal to the sum of $\Delta E_{OH\ 1:1}$ and $\Delta E_{OH\ 1:2}$ in the case of higher degree of association. The knowledge of $\Delta E_{OH\ 1:2}$ allows to assume the concentration of 1:2 associates and to calculate the equilibrium constant $K_{1:2}$.

In order to get an impression on the magnitude of the equilibrium constants to expect under our experimental conditions n-hexanol was investigated regarding its association behaviour with pyridines. The results are presented in table 45.

Table 45: Equilibrium constants $K_{1:1}$ for the intermolecular association of n-hexanol with pyridine and its derivatives in methylene chloride at 12 and 32 °C

	K [in l/mol]		
	Pyridine	2.4-Lutidine	2.4.6-Collidine
at 12 °C	1.9	1.7	3.4
at 32 °C	1.4	1.0	2.3

The equilibrium constants are somewhat unexpected concerning the sequence of the equilibrium constants for pyridine and 2.4-lutidine, but they may characterise the order of magnitude for mono aclohols.

In diols with no substantial intramolecular hydrogen bonding like hexan-1.6-diol, pentan-1.5-diol and cyclohexan-1.4-diol the equilibrium constants $K_{1:1}$ can be calculated analogously if the amine concentration was lower than 0.1 mol/l (Table 46).

Table 46: Equilibrium constants $K_{1:1}$ for diols with amine derivatives in methylene chloride

		Cyclohexan-1.4-diol			Pentan-1.5-diol			Hexan-1.6-diol		
		Py	Lu	Col	Py	Lu	Col	Py	Lu	Col
at 12° C	$K_{1:1}$ [in l/mol]	5.6	7.8	6.0	2.1	6.5	6.5	4.0	7.5	4.5
	$K_{1:2}$ [in l^2/mol^2]	2.7	3.6	2.5	1.7	4.2	0.0	4.2	4.6	4.0
at 32 °C°	$K_{1:1}$ [in l/mol]	3.5	5.0	4.5	0.7	4.5	4.0	3.5	4.2	2.8
	$K_{1:2}$ [in l^2/mol^2]	2.1	2.0	0.9	0.8	1.6	0.0	2.5	2.0	0.0

Py=pyridine, Lu=2.4-lutidine, Col=2.4.6-collidine

Now we turn our interest to alcohols in which intramolecular interactions cannot be suppressed under our conditions. Rather strong intramolecular hydrogen bonds are presented in the spectra of propan-1.3-diol as well as of butan-1.4-diol with OH bands shifted down up to 80 cm^{-1} and 188 cm^{-1}, respectively. But the intramolecular hydrogen bond is evidently not extended over all molecules; for in the solution of butan-1.4-diol the integral molar extinction coefficient of the free band near 3618 cm^{-1} is ca. 40% higher than that of the intramolecular associate. Judging from our experience the intensity ratio mean that about 30% of the molecules are intramolecularly hydrogen bonded and they are exhibiting a remarkable temperature effect on the intensities of the associate bands. But this effect should not be used to estimate the hydrogen bond enthalpy because the associate OH band is much more temperature sensitive than the free OH band. Compared to the intensities of the free OH band in hexan-1.6-diol and pentan-1.5-diol the extinction of the free OH band is rather high and this may be perceived as a hint for a strong co-operative effect originated by the intramolecular hydrogen bonding. Hence, strictly speaking we have to expect two bands at the position of the free band namely the first one of the really free band and the second one of that OH bond which is acting as an acceptor in the intramolecular bond as it was sketched in Fig. 47 and 48. Although by no means one may reduce an overlap of two free bands at this position there should be two bands with different extinction coefficients causing problems in the calculation.

The equilibrium constants in the systems with butan-1.4-diol are extremely large compared with those using pentan-1.5-diol as proton donor. Already at low amine concentrations of ca. 0.1 mol/l more than half of the intensity of the free band vanishes in favour of an associate band. This means that the association need to be discussed in several steps. Firstly, the free OH group of the intramolecularly bonded diol should be attacked because of its high polarity. As a consequence of the decrease in the concentration of free OH functions the intramolecular hydrogen bond equilibrium will produce new co-operativity affected OH groups which can be again a centre for the acceptor attack.

On the other hand a large amount of the alcohol molecules should offer both OH groups for the association. In this case the co-operativity argument should not be valid. At a high degree of association there should be a decrease in the co-operativity aided intermolecular bonds.

In table 47 there are given the equilibrium constants K for the 1:1 and 1:2 associate systems of butan-1.4-diol and tertiary amines at 12 °C and 32 °C.

Table 47: Equilibrium constants for the association of butan-1.4-diol with pyridine and its derivatives

		Butan-1.4-diol		
		Pyridine	2.4-Lutidine	2.4.6-Collidine
at 12 °C	$K_{1:1}$ [in l/mol]	32	56	90
	$K_{1:2}$ [in l^2/mol^2]	4.7	8.6	22
at 32 °C	$K_{1:1}$ [in l/mol]	26	54	87
	$K_{1:2}$ [in l^2/mol^2]	4.4	8.3	19

The large difference in $K_{1:1}$ between the differently basic amines and butan-1.4-diol substantiates the importance of the high polarity of the OH groups in these systems. As it was expected $K_{1:2}$ is evidently smaller than $K_{1:1}$.

For propan-1.3-diol the problems originated by co-operativity described for butan-1.4-diol will principally also appear in this compound. Also the spectrum of this molecule gives evidence for an internal hydrogen bond which is realised by a six-membered ring. Obviously this intramolecular hydrogen bond is weaker than in butan-1.4-diol, for $\Delta\nu_{OH}$ is significantly lower as well as the intensity of the associate band. The equilibrium constants are given in table 48 for amine concentrations of ca. 0.05 mol/l.

Table 48: Equilibrium constants for the association of propan-1.3-diol with pyridine and its derivatives in methylene chloride

			Propan-1.3-diol	
		Pyridine	2.4-Lutidine	2.4.6-Collidine
at 12°C	$K_{1:1}$ [in l/mol]	12	12	12
	$K_{1:2}$ [in l^2/mol^2]		25	26
at 32° C	$K_{1:1}$ [in l/mol]	8	9	13
	$K_{1:2}$ [in l^2/mol^2]		7	8

This means that in propan-1.3-diol the co-operativity effect is less dominant than in butan-1.4-diol. The seven membered ring in that molecule evidently allows to pass on the co-operativity influence than the six membered ring of propan-1.3-diol but the co-operative effect is acting here still and explains the increase of the equilibrium constants in comparison to hexan-1.6-diol and pentan-1.5-diol.

Another fact should be mentioned here: Whereas at concentrations of amine lower than 0.1 mol/l the apparent equilibrium constant usually remained constant, in propan-1.3-diol already at an amine concentration higher than 0.05 mol/l a considerable amount of 1:2 associates occur (Table 49). This interpretation is supported by the fact that the absorptivity of the free band in both molecules is ca. 15% higher than in hexan-1.6-diol or pentan-1.5-diol.

Table 49: Influence of the amine concentration of the apparent equilibrium constant K_{app} according to eq. 16 for the associate formation between propan-1.3-diol and collidine (Density changes due to the different temperatures have been considered in the calculation.)

$c_{alcohol}$ [in mol/l]	$c_{collidine}$ [in mol/l]	K_{app} at 12 °C	K_{app} at 32 °C
0.02934	0.0595	12.2	11.8
0.02934	0.131	27.3	16.9
0.02934	0.1587	38.5	15.9
0.02934	0.1723	56.9	19.3

In diols with the OH functions in *ortho* position like propan-1.2-diol, *cis*-cyclohexan-1.2-diol and *trans*-cyclohexan-1.2-diol the equilibrium constants in table 43 have shown that these molecules are fully intramolecularly hydrogen bonded. Thus, the separated OH signals of these diols can be taken as measure for the concentration of both OH species. When discussing the intermolecular association we have to agree that the intramolecular equilibrium is not affected. We can substantiate applicability of this approach since we did not observe any symmetry changes of the OH band. Furthermore, *ab initio* calculations on propan-1.2-diol have shown that intramolecular associates are essentially retained in the presence of a water

molecule which either acts as proton donor or acceptor with one diol molecule[240]. This observation is in strong contrast with that of propan-1.3-diol where the intramolecular hydrogen bond is broken in favour to the intermolecular associate.

The main result for these diols is that both OH groups of one molecule act practically independently of the neighbour OH. This is valid for a large range of amine concentration where the equilibrium constants remain constant as it is shown in table 50 for the equilibrium between *cis*-cyclohexan-1.2-diol and 2.4-lutidine at 12 °C.

Table 50: Equilibrium constant for the association between the two OH species of *cis*-cyclohexan-1.2-diol and 2.4-lutidine at 12 °C

c_{amine} [in mol/l]	0.0657	0.1235	0.1600	0.2165	0.2765	0.3725	0.8325
K [in l/mol] (at 3611 cm^{-1})	6.41	5.38	5.64	6.69	6.87	6.55	7.43
K [in l/mol] (at 3577 cm^{-1})	1.53	1.79	1.52	1.94	1.63	1.87	1.78

The data in table 51 represent the 1:1 associates of the OH function which is activated by the intramolecular interaction

Table 51: Equilibrium constants for the intermolecular association of selected diols with pyridine and its derivatives in methylene chloride

$K_{1:1}$ [in l/mol]		Propan-1.2-diol		
	T in °C	Pyridine	2.4-Lutidine	2.4.6-Collidine
at 3620 cm^{-1}	12	9.8	10	22.0
	32	8.5	7.5	18.0
at 3588 cm^{-1}	12	2.3	3.1	3.6
	32	3.3	3.5	4.0
			cis-Cyclohexan-1.2-diol	
at 3611 cm^{-1}	12	3.6	6.5	7.8
	32	2.4	5.0	4.3
at 3577 cm^{-1}	12	1.2	1.7	1.2
	32	1.4	1.2	1.0
			trans-Cyclohexan-1.2-diol	
at 3613 cm^{-1}	12	11.8	21.0	16.4
	32	8.4	18.0	14.3
at 3584 cm^{-1}	12	2.0	4.4	3.7
	32	2.0	2.8	3.1

Furthermore, we can see that the association is always stronger for the OH function in the diol which acts as proton acceptor in the intramolecular associate. As it was observed for the association of monomeric alcohols and diols with no intramolecular associates the increasing basicity of the pyridine derivatives causes generally higher equilibrium constants. Deviations from this general rule might be caused by steric effects. Comparing the association behaviour of propan-1.3-diol and propan-1.2-diol the equilibrium constants with pyridine are slightly larger in propan-1.3-diol substantiating the co-operativity effect discussed above on the example of butan-1.4-diol.

The *cis/trans* isomers of cyclohexan-1.2-diol are influenced in their association by the arrangement of the OH function in analogy to the effects discussed on a intramolecular association. The diequatorial oriented OH groups are significantly superior in the intermolecular association than the equatorial-axial geometry of the OH where the co-operativity cannot be as effective.

We have attempted to describe the relations between the equilibrium constants $K_{1:1}$ and $K_{1:2}$ in a statistical manner. For diols with OH groups which do not influence each other in their hydrogen bond forming affinity which is the case if the distance between both functions is large like in hexan-1.6-diol, pentan-1.5-diol or in cyclohexan-1.4-diol the addition of amine will produce three different diol species: alcohol molecules which can be attacked by one (1:1) or two (1:2) amine molecules or not being attacked at all. The snap shot in the solution will be comparable with a geometrical model namely a field partitioned in equally large areas (diols) which may catch 0 or 1 or 2 other particles but not more. The number of areas should be N with 2N places. The number of amines should be n. Now the particles (amines) will be spread on the field and the situation will be as follows: Some may fail to interact at all (free amine) but the others will be placed on the areas. Learning from statistics, the probability to catch one amine is n_1/N, if n_1 is the number of amines we used in a first step. But this consideration is not conclusive because there is also the possibility, that some areas may possess two amine molecules. This chance is given by the square of the simple probability. Then the equation we use is the following:

$$\frac{n_1}{N} = \frac{2}{x^2} + \frac{1}{x} \qquad [19]$$

Herein the factor 2 is the number of molecules in double occupied fields. In the case of N=600 and n_1=100 the probability x is equal to 7.6. This means that in our case the probability to receive 1:1 associates is ca. $\left(\frac{1}{7.6} \div \frac{1}{6}\right) \cdot 100 = \frac{6}{7.6} \cdot 100 \approx 79$ but $\frac{6}{7.6^2} \cdot 100 \approx 10$ areas may be already possess two amine molecules. The respective values for n=200 are 137 areas with one amine molecule and ca. 31 bearing two amine molecules (for n= 400 results 227 and 86, respectively). In the case of n=N equi-distribution results, this means 300 single and 150 double occupied areas will be found and still 150 remain empty. If n becomes larger than N the consideration are ambiguously different: Now we start from the situation that all areas should be engaged twice and we remove some particles (amines) statistically. This is a symmetrical procedure as we have described above.

As a conclusion we learn that statistically the number of single occupied areas do not exceed the number of 0.5N. With n growing larger than N this number decreases and approaches zero in the case of a very high amine concentration (Fig. 51)

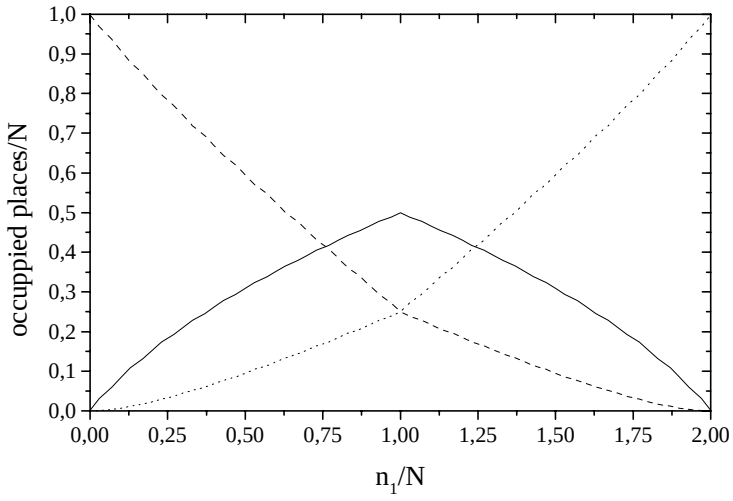


Fig. 51: Statistical occurrence of single (solid line) and double (dotted line) occupied and free (dashed line) places in dependence of n_1/N

This consideration allows to describe the relation of $K_{1:1}$ and $K_{2:1}$ for diol-amine associates. Starting from the law of mass in the case of 1:1 association $K_{1:1}$ is given by

$$K_{1:1} \cdot c_B = \frac{c_{ass1:1}}{c_A} \qquad [20]$$

and

$$K_{1:2} \cdot c_B^2 = \frac{c_{ass1:2}}{c_A} \qquad [21]$$

From the first equation it follows, that

$$c_A = \frac{c_{ass1:1}}{K_{1:1} \cdot c_B} \qquad [22]$$

Substituting this expression in eq. 20 we receive

$$K_{1:2} \cdot c_B^2 = \frac{c_{ass1:2} \cdot K_{1:1} \cdot c_B}{c_{ass1:1}} \qquad [23]$$

and by choosing the experimental conditions $c_{B0} >> c_{A0}$, the acceptor concentration is $c_B = c_{B0}$

$$K_{1:2} = \frac{c_{ass1:2}}{c_{ass1:1}} \cdot \frac{K_{1:1}}{c_{B0}} \qquad [24]$$

Following from our discussion the statistical relation between $K_{1:1}$ and $K_{1:2}$ is exactly 1/2, when just half of the OH groups are associated. In the infrared spectra we would refer to an OH band being half of its initial intensity representing c_{A0}. Our procedure now is to determine from our experimental results the concentration of amine just decreasing the intensity of the free OH band to 0.5. This will be done for the diols having equivalent OH groups namely hexan-1.6-diol, pentan-1.5-diol and cyclohexan-1.4-diol in association with the pyridine derivatives. The experimental conditions were chosen according to the statistical requirements which means that the OH intensity after the addition of the amine decreases to half of its initial intensity. To give an example, for a 0.05 mol/l hexan-1.6-diol solution 0.6 mol/l pyridine need to be added to fulfil the requirements.

The equation which will describe $K_{1:2}$ is

$$K_{1:2} = \frac{1}{2} \cdot \frac{K_{1:1}}{c_{B0}} = \frac{4.2}{2 \cdot 0.6} = 3.8 \qquad [25]$$

The experimental value for $K_{1:2}$ was 2.5 under the experimental conditions mentioned above. In the systems of pentan-1.5-diol with pyridine, 2.4-lutidine and 2.4.6-collidine we calculate based on statistical distribution 3, 6.5 and 6 as values of $K_{1:2}$ respectively, but experimentally we have found in these cases 1.7, 4.2 and 0.0. Also for cyclohexan-1.4-diol the

experimental values of $K_{1.2}$ are lower than those deduced by statistical consideration (Table 52).

Table 52: Comparison of the experimental and statistical $K_{1:2}$ values

$K_{1:2}$	Hexan-1.6-diol	Pentan-1.5-diol			Cyclohexan-1.4-diol		
	Py (32 °C)	Py (12 °C)	Lu (12 °C)	Col	Py (at 32°C)	Py (at 32 °C)	Lu (at 32°C)
experimental	2.5	1.7	4.2	0.0	2.1	2.7	2.0
statistical	3.8	3.0	6.5	6.0	3.0	4.75	3.75

As the conclusion from these data it can be seen that also in the systems with some distance between the OH groups these groups are not independent from each other, but the interaction between the acceptor molecules which could be sterically or electronically reasoned may cause the differences between the values statistically expected and measured values. This interaction always decreases the possibility to form 1:2 hydrogen bonds.

B.3.4.2 Intramolecular association behaviour of peptide derivatives

In paragraph B.3.2.2.5 we have discussed the NH and amide I band profiles of some tripeptide derivatives dissolved in carbon tetrachloride. Furthermore, the occurrence of the broad NH associate bands were assigned to intramolecular hydrogen bonding forming five or seven membered associate rings. The occurrence of these hydrogen bonded species is dependent on the amino acid sequence. Thus, whereas the diastereoisomers of Z-Ala-Phe-Val-OMe and Z-Ala-Leu-Val-OMe form a variety of intramolecular hydrogen bonds, Z-Phe-Ala-Val-OMe has only one associate signal at low temperatures. That means, only a small proportion of the available NH groups form C_7 associates in carbon tetrachloride solution. Contrary the NH band positions and the integral molar extinction coefficients of the tripeptides dissolved in methylene chloride are independent of the sequence and chirality of the amino acid units and rather similar in magnitude of ε_{Mol} (Table A13 and A14). Even the substitution of Phe by another hydrophobic amino acid results in only small changes.

Hence we can establish an approach for the calculation of the intramolecular association constant similar to the one shown for diols which is based on the correlation of the intensity of the free NH bands of different molecules.

Therefore, the following approximation should be made:

- In carbon tetrachloride solution of Z-Phe-Ala-Val-OMe the intensity of the free NH signal was very weak. Therefore, nearly the whole peptide concentration is represented by the C_5 NH signal.
- Neglecting the small degree of association of Z-Phe-Ala-Val-OMe at lower temperatures the integral of the C_5 signal can be used for the calculation the equilibrium constants K via the decrease of this band.
- The temperature dependency of the molar extinction coefficients of the signals are similar.

The results of this approximation are summarised in table 53.

Table 53: Equilibrium constants K for the formation of C_7 rings of tripeptide derivatives dissolved in carbon tetrachloride in dependence of the temperature

Substances	K at				
	-10 °C	-5 °C	15 °C	25 °C	40 °C
Z-Ala-Phe-Val-OMe					
LLL	2.95	2.29	1.04	0.65	0.39
LDL	2.31	1.90	1.09	0.85	0.63
DLL	3.89	2.75	1.15	0.77	0.51
Z-Ala-Phe-Val-OtBu	2.60	2.27	0.69	0.44	0.22
Z-Ala-Leu-Val-OMe					
LLL	3.05	2.42	0.98	0.63	0.39
LDL	2.23	1.89	1.16	0.95	0.74

Comparing the data of the LLL and LDL diastereoisomers of Z-Ala-Phe-Val-OMe and Z-Ala-Leu-Val-OMe at -10 °C the LLL diastereoisomer associates more strongly than the LDL analogue. Between LLL and DLL diastereoisomers approximately the same values were calculated.

It is interesting to note that latest investigations on the mechanism of trans activation of HIV-1 gene expression have identified a small family of mostly heterochiral tripeptides which are capable of structure-specific binding to the bulge loop of the TAR RNA[243]. *In vitro* binding studies of tripeptide with polar side functions give binding constants between 1.73 and 0.07.

The temperature dependence of K can be graphically expressed as a plot of lnK vs. 1/T (Fig. 52) or the enthalpy values $\Delta_R H$ can be calculated according to the Van't Hoff relation (Table 54).

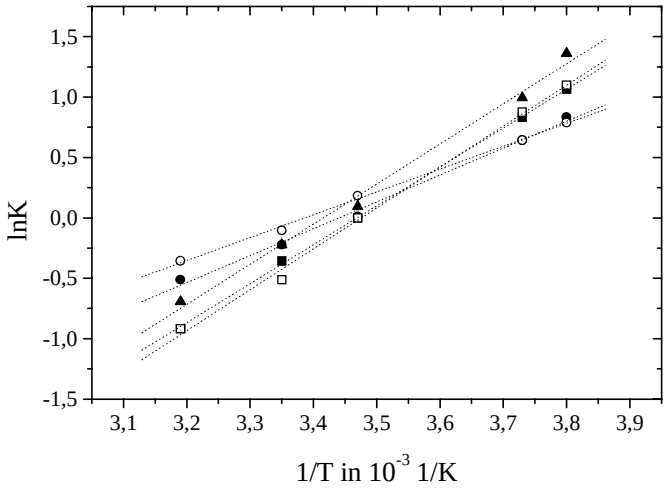


Fig. 52: ln K - 1/T- plot for the diasteroisomers of Z-Ala-Phe-Val-OMe (solid) and Z-Ala-Leu-Val-OMe (open symbols, square: LLL, circle: LDL, triangle: DLL)

The deviation from linearity in the considered temperature interval are rather small. Hence the linear regression gives a correlation of 0.99. We had expected a larger deviation because the equilibrium constants represent the summation over different hydrogen bonds.

Table 54: $\Delta_R H$ values of the association of C_7 rings in tripeptide derivatives dissolved in carbon tetrachloride [in kJ/mol]

Substances	$\Delta_R H$
Z-Ala-Phe-Val-OMe	
LLL	-27.8
LDL	-17.7
DLL	-27.9
Z-Ala-Phe-Val-OtBu	-33.5
Z-Ala-Leu-Val-OMe	
LLL	-28.1
LDL	-15.0

The Badger-Bauer rule[244] describes the correlation between the enthalpy values and the shift of the free NH to the maximum of the associate NH. In the case of the tripeptides discussed we have observed two associate bands in the NH fundamental region. The intense

signal shifted ca. 105 to 110 cm^{-1}, which corresponds to the C_5 arrangement, cannot explain the considerable differences in the enthalpy values, although the signal behaves very sensitively to temperature changes. The differences are rather small, as may be seen by comparing the associate band shift of 58 cm^{-1} for the LDL diastereoisomer and 45 cm^{-1} for the LLL and DLL compounds, respectively. The position of these bands is slightly uncertain, due to the strong overlapping of the signals. So we tend to explain the differences more in terms of steric effects than in electronic.

B.3.4.2.1 Effects of intramolecular association on the diffusion behaviour of tripeptide derivatives

The migration of a molecule through biological membranes depends on its hydrophobicity and hence on the combination of numerous interactions. Hydrogen bonding and dipole interactions are among the relevant interactions which will occur when passing a membrane or simply diffusion in a media as obeyed in the 1. Fick`s law (eq. 26).

$$v = -\frac{k_B T}{6\pi\eta r} \cdot \frac{d \ln c}{dx} \qquad [26]$$

Herein the diffusion rate depends strongly on osmotic forces. Furthermore the viscosity of the medium η and the radius of the particles influences the migration.

Mechanistic studies on the influence of intermolecular hydrogen bonding on the diffusion behaviour have been performed using alcohol/amine systems in different solvents and different temperatures[245, 246, 247]. It was found that the intermolecular hydrogen bonding between the benzyl alcohol and proton acceptor increases the diffusion rate. A clear dependence between the strength of the formed associate represented by the equilibrium constant and the diffusion rate could be stated from these investigations.

Table 55: Diffusion rates v of the tripeptide derivatives in carbon tetrachloride and toluene [in 10^{-2} mol/l*min]

Substance	v in CCl_4	v in Toluene
Z-Ala-Phe-Val-OMe		
LLL	9.96	780
LDL	17.70	960
DLL	8.64	720
Z-Ala-Leu-Val-OMe		
LLL	19.62	2057
LDL	16.32	1320
DLL	17.76	1328

In this study it was our intention to prove if the intramolecular association of tripeptide derivatives affects the diffusion behaviour. Therefore, the investigations were performed in carbon tetrachloride and toluene. As we know from the considerations above intramolecular hydrogen bonds are observed in the former solvent whereas they will break apart in the more polar toluene. The diffusion rates v are given in table 55.

According to eq. 26 the particle size of the molecules and the dynamic viscosity of the solutions will determine the diffusion behaviour. The particle size of the non-solvated molecule may be expressed by the characteristic McGowan volume V_x which was calculated 3.776 and 3.591 10^{-2} l/mol, for the diastereoisomers of Z-Ala-Phe-Val-OMe and Z-Ala-Leu-Val-OMe, respectively. In agreement with this relation we find that the less voluminous tripeptide derivatives, Z-Ala-Leu-Val-OMe, migrate faster than Z-Ala-Phe-Val-OMe in both solvents.

Regarding the viscosity of the solutions a ratio of the dynamic viscosity of toluene and carbon tetrachloride was estimated to be 0.61[247]. Thus, according to the theory the diffusion rate of the tripeptide derivatives should be 1.5 to 2 times faster in toluene. However, the experimental finding of the diffusion rates in toluene and in carbon tetrachloride are astonishing. In toluene the diffusion rates are accelerated 50 to 100 times. Obviously, viscosity changes are not sufficient to explain the migration enhancement in toluene. If the proton acceptor and the polar properties of both solvents are compared the descriptors of the UCL database might be helpful. Due to its conjugated electron system toluene exhibits proton acceptor properties (($\sum\beta_2^H$= 0.14) whereas the effective hydrogen bond basicity of carbon tetrachloride is equal to zero. Furthermore, the polar properties (π_2^H) are 0.38 and 0.52, for carbon tetrachloride and toluene, respectively.

Due to the differences in the solvent characteristics toluene is able to destroy the intramolecular hydrogen bonds in the tripeptide derivatives by forming even stronger intermolecular hydrogen bonds between the NH functions and the phenyl ring. As it was observed in the solid state this type of interaction do occur in peptide derivatives[248].

With regards to the effect of intramolecular hydrogen bonding in diastereoisomers a higher diffusion rate was found for the LDL form of Z-Ala-Phe-Val-OMe than for the LLL and DLL diastereoisomers in carbon tetrachloride. This corresponds well with the equilibrium constants K which are a measure for the stability of C_7 associates at 25 °C (table 53). In the case of Z-Ala-Leu-Val-OMe the correlation between K and the diffusion rate is not as obvious.

B.3.4.3 Intermolecular association behaviour of amino acid and peptide derivatives with alcohols in methylene chloride

B.3.4.3.1 Effect of the addition of 1.1.1.3.3.3-hexafluoropropanol to peptide solutions

The intention of the IR studies was the determination of equilibrium constants between proton donor molecules and the oligopeptide derivatives. It was shown in the former paragraphs that logK for a large number of 1:1 hydrogen bond associates correlate well with the hydrogen bond basicity β_2^H when carbon tetrachloride is used as a solvent. The low solubility and the strong intramolecular association behaviour of our compounds in carbon tetrachloride leads us to prove if a correlation between logK of 1:1 associates and β_2^H exists in methylene chloride as well. Indeed, this was found within a group of molecules with same acceptor functions. Thus, associate complexes of alcohol molecules and proton acceptors of the structure X=O (X=P, S, C, As) and C-Y-C (Y=O, S) give a good correlation.

Oligopeptides are relatively weak proton acceptor compounds. Hence in order to receive accurate equilibrium constants strong proton donors such as HFiP are required. When studying the association behaviour of fluorinated alcohols with oligopeptides some peculiarities have to be kept in mind:

Fluorinated alcohols are capable to change the secondary structure of peptides and proteins. Hence, especially at low concentrations a more ordered arrangement in the peptide backbone can be observed due to the helicogenic effect[249, 250]. Di- and tripeptide derivatives occur in an unordered arrangement in methylene chloride and in HFiP. However, as the amide I band profiles indicated the population of the stable conformers is different due to the properties of both solvents. The selection of the measurement conditions should be done under two aspects. Firstly, no intermolecular association of the alcohol molecules should occur. Secondly, an

association on the acceptor compounds should be observed but changes of the conformational equilibrium should be avoided because this would affect the band profile in the amide I region.

Further problems are justified by the nature of oligopeptides. In contrast to acceptor compounds such as acetamide normally used for the investigation of equilibrium constants a tripeptide derivative contains at least four C=O acceptor functions with similar hydrogen bond basicity. Regarding the accessibility of the acceptor functions differences are expected in dependence on the bulkiness of the amino acid residues and the location of the function within the amino acid sequence.

Combined NMR and IR studies would allow to locate the positions of the alcohol attack.

B.3.4.3.1.1 ^{1}H NMR titration

Selected peptide solutions were studied by ^{1}H NMR spectroscopy in order to identify the preferred amide functions in the molecule for an attack of the proton donor. In the preliminary step of NMR titration the routine spectra in $CDCl_3$ solution were recorded and with the help of the ^{1}H, ^{1}H COSY spectra the main signals were assigned (Table 56).

Summarising the data in table 56 the ^{1}H NMR the NH protons of the urethane functions are always found at higher fields than the peptide NH signals. Furthermore, the chemical shift of the peptide protons are determined by the type of the amino acid, its position in the peptide backbone and its chirality. For substances with no defined secondary structure the proton chemical shift of NH and $C^{\alpha}H$ will represent the average distribution of the variety of conformers.

Table 56: Chemical shift [in ppm] of significant signals in the ^{1}H NMR spectra of selected compounds dissolved in $CDCl_3$. Assignment is based on ^{1}H, ^{1}H COSY spectra.

Substance	$CH_2OC(O)$	NH	$C^{\alpha}H$	$C^{\beta}H$	OCH_3
Z-^{1}Ala-^{2}Ala-OMe	5.142				3.73
^{1}Ala		5.26	4.223	1.374	
^{2}Ala		6.429	4.561	1.374	
Z-^{1}Ala-^{2}Ala-^{3}Ala-OMe	5.142				3.727
^{1}Ala		5.265	4.215		
^{2}Ala		6.594	4.447		
^{3}Ala		6.554	4.523		
Z-^{1}Ala-^{2}Ala-^{3}Val-OMe	5.143				3.718
^{1}Ala		5.259	4.221	1.365	
^{2}Ala		6.546	4.467	1.365	
^{3}Val		6.509	4.479	2.152	
Z-^{1}Ala-^{2}Val-^{3}Val-OMe	5.140				3.714
^{1}Ala		5.308	4.241	1.359	
^{2}Val		6.567	4.241		
^{3}Val		6.377	4.501		
Z-^{1}Val-^{2}Val-^{3}Val-OMe	5.087			n.d.	3.710
^{1}Val		5.396	4.002		
^{2}Val		6.4545	4.261		
^{3}Val		6.4545	4.506		
Z-^{1}Val-^{2}D-Val-^{3}Val-OMe	5.079				3.681
^{1}Val		5.3435	3.9835	n. d.	
^{2}D-Val		6.474	4.3395		
^{3}Val		6.515	4.488		

n.d.= not detectable; the signals are strongly overlapped and cannot be assigned to the different amino acid residues.

The addition of alcohol to an apolar peptide solution will result in changes of the molecular environment due to the proton attack on the acceptor function of the amide group and most probably to conformation variation which can be monitored sensitively on the shift of the amide protons[251]. In constraint conformers the relative shift of the NH band can be used for the semi-quantitative description of the equilibrium constant for the hydrogen bonded associate complex[252.] For oligopeptides the results of the ^{1}H NMR titration must be discussed more complex. While hydrogen bonding would generally lead to a down field shift of the NH protons the change of the conformation cannot be predicted as easily. However, from the comparison of the responses of the NH signals on the attack of HFiP it is possible to conclude on the accessibility of the amide groups within the molecule assuming similar hydrogen bond

basicities of the acceptors. In Fig. 53 the relative chemical shifts of the NH and $C^{\alpha}H$ signals of Z-^{1}Ala-^{2}Ala-OMe vs. the volume of HFiP added to 1 ml of the peptide solution.

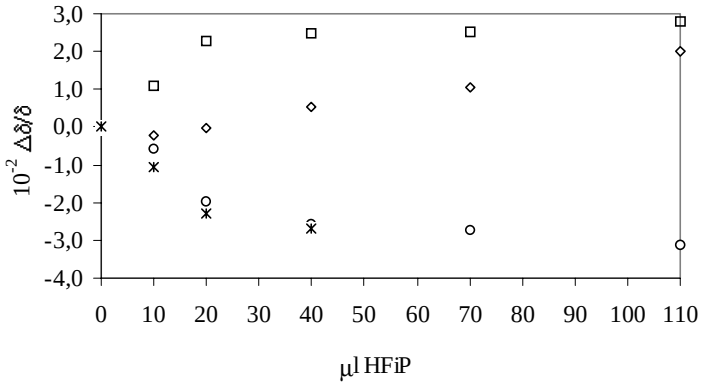


Fig. 53: Relative chemical shift changes (Δδ/δ) of ^{1}H NMR signals of Z-^{1}Ala-^{2}Ala-OMe with the addition of HFiP (NH: ^{1}Ala: triangle, ^{2}Ala: square, $C^{\alpha}H$: ^{1}Ala: cross, ^{2}Ala: circle)

Both NH signals shift to higher frequencies but the shift of the urethane amide function is delayed regarding to the amide signal. The behaviour of the $C^{\alpha}H$ signals substantiates the observed differentiation of the functions but generally a shift to higher fields is observed. Although the amide function of ^{2}Ala is located relatively in the medium region of the molecule it seems easily accessible whereas the NH linked to the benzyloxycarbonyl protecting function is blocked for the HFiP attack. Thus, conformational changes induced by the changed solvent polarity must occur in prior to effective association.

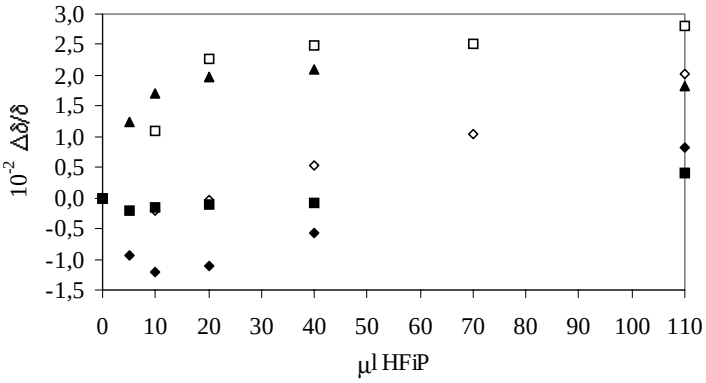


Fig. 54: Relative chemical shift changes (Δδ/δ) of the NH protons of Z-^{1}Ala-^{2}Ala-OMe (open) and Z-^{1}Ala-^{2}Ala-^{3}Ala-OMe (solid) with the addition of HFiP (^{1}Ala: diamond, ^{2}Ala: square, ^{3}Ala: triangle)

In methylene chloride solution of Z-^{1}Ala-^{2}Ala-^{3}Ala-OMe the preferred acceptor position is the third amino acid which shows even at low alcohol concentrations a significant shift of the

NH signal (Fig. 54). Contrary respond of the ^{2}NH is rather low whereas the ^{1}NH (urethane) group shows a shift to higher fields demonstrating that this function becomes more shielded at low alcohol concentrations. Only at high alcohol amounts (110 µl) $\Delta\delta/\delta$ receives a positive sign as expected by the formation of hydrogen bonds.

In conclusion of these studies it can be stated that the conformational equilibrium present in methylene chloride which is determined by the polar interaction between solvent and solute does not support the formation of intermolecular hydrogen bonds with HFiP. The interaction of HFiP with the peptide seems to start at the most outer amide function not protected by the bulky Z residue. As a consequence of this interaction the peptide backbone may adopt a conformation which causes a larger shielding of the other NH functions. At higher amounts of HFiP the polar properties of the alcohol will force a conformation in which all acceptor positions can interact in perfect geometry with the alcohol.

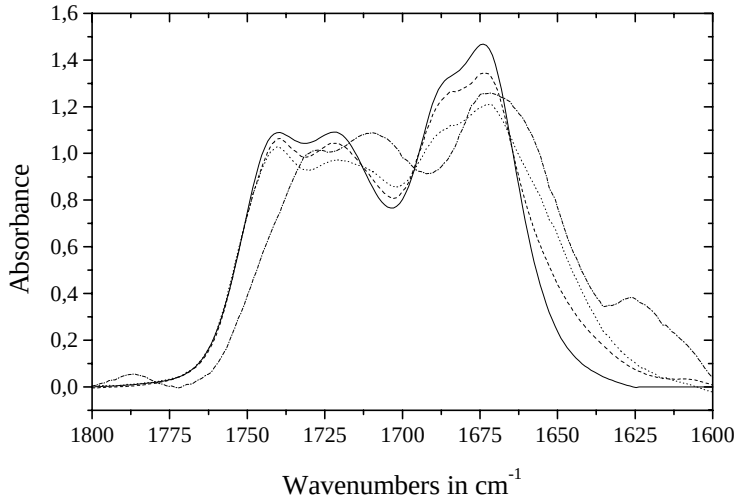


Fig. 55: Carbonyl and amide I region of Z-Ala-Ala-Ala-OMe dissolved in methylene chloride and in addition of defined amounts of HFiP (solid line: pure peptide solution, dashed: 1 ml solution+ca. 1µl HFiP, dotted: 1 ml solution+5 µl HFiP, dash-dotted: 1 ml solution+110 µl HFiP; d=3 mm NaCl)

In Fig. 55 the analogue alcohol titration was monitored in the amide I region of the mid infrared spectra which is more sensitive than the NMR spectroscopy. Already with ca. 1 µl of HFiP the band profile indicates the occurrence of associate bands located at 1710 and 1651 cm^{-1}. It should be mentioned here that at low alcohol concentrations the ester carbonyl function does neither change in its intensity nor a new signal could be observed which would indicate a hydrogen bond formation at this acceptor function in the molecule. From steric

reasons this could not be understood. It can be only explained by its lower basicity which makes the ester function a non-attractive acceptor group.

With an amount of 110 µl HFiP per 1 ml solution the spectra almost exhibits the contours of the spectra recorded in solution of pure HFiP. Hence the solvent-solute interactions are clearly determined by the alcohol and not by the apolar methylene chloride.

Generally similar behaviour as for Z-^{1}Ala-^{2}Ala-^{3}Ala-OMe is observed for the diastereoisomers of Z-^{1}Val-^{2}Val-^{3}Val-OMe in the ^{1}H NMR titration (Fig. 56) although the accessibility of the amide functions on ^{2}Val and ^{3}Val is not as much hindered as in the case of Z-Ala-Ala-Ala-OMe. Conclusively the distribution of the conformers supports a proton attack which would be the case in the extended conformation of the peptide chain. The bulky residues will be the reason for the higher amount of extended conformers.

Regarding the influence of the diastereoisomerism the accessibility of the peptide function linking the second and the third amino acid is favoured for the interaction with HFiP whereas the behaviour of the peptide group ^{1}Val-^{2}Val is comparable in both diastereoisomers.

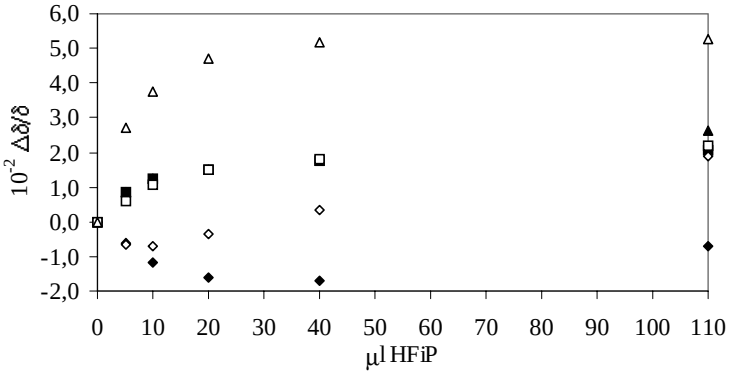


Fig. 56: Relative chemical shift changes (Δδ/δ) of the NH protons of Z-^{1}Val-^{2}Val-^{3}Val-OMe (solid) and Z-^{1}Val-D-^{2}Val-^{3}Val-OMe (open) with the addition of HFiP (^{1}Val: diamond, ^{2}Val: square, ^{3}Val: triangle)

For the determination of the equilibrium constant for the association of HFiP and oligopeptide derivatives the concentration of HFiP should be small in order to avoid conformational changes in the peptide backbone. As CD and NMR spectra show the conditions are fulfilled at alcohol concentrations of 0.005 to 0.01 mol/l.

B.3.4.3.1.2 Complexity of the calculation of the equilibrium constants due to the presence of multiple acceptor positions in peptide derivatives

Contrary to the alcohol systems where the proton donor properties were studied in the peptide derivatives the hydrogen bond basicity is of interest. Hence the decrease of the OH signal due to the addition of the peptide derivative will give an equilibrium constant representing the overall basicity of the molecules.

The difficulty in the calculation lies in the occurrence of a number of acceptor functions which are not equal in their hydrogen bond basicity and as we have concluded from the ^{1}H NMR titration are not equally accessible. This, however, also implies a challenge for the calculation of individual K values since the acceptor functions show separate signals in the carbonyl stretching region. These signals can be separated by band deconvolution and peak fit programs and used to monitor the associate formation with HFiP by the decrease of the corresponding acceptor signal.

For a dipeptide derivatives B with at least three potential non-equivalent acceptor functions a, b and c (urethane CO, peptide CO, ester CO) at low HFiP concentration the alcohol can interact with each function giving equilibrium constants K_a, K_b and K_c.

$$B^aB^bB^c + A \xrightarrow{K_a} B^a(A)B^bB^c \qquad [27]$$

$$B^aB^bB^c + A \xrightarrow{K_b} B^aB^b(A)B^c \qquad [28]$$

$$B^aB^bB^c + A \xrightarrow{Kc} B^aB^bB^c(A) \qquad [29]$$

Thus, for the calculation of the equilibrium constants on each acceptor function the accessible (free) alcohol concentration is reduced by the sum of associate concentrations on a, b and c.

The eq. 16 applied on the acceptor function a can be written as follows.

$$K_a = \frac{c^a_{ass}}{(c_{A0} - c^a_{ass} - c^b_{ass} - c^c_{ass}) \cdot c^a_B} \qquad [30]$$

In an analogue approach K_b and K_c can be calculated. The overall equilibrium constant for a substance with three acceptor functions can be calculated as

$$K_{all} = \frac{c^{a}_{ass} + c^{b}_{ass} + c^{c}_{ass}}{(c_{A0} - c^{a}_{ass} - c^{b}_{ass} - c^{c}_{ass}) \cdot (c^{a}_{B} + c^{b}_{B} + c^{c}_{B})} \qquad [31]$$

For our calculation we have used the integral intensities of the deconvoluted signals in the amide I region which have been assigned in the former paragraph. For most accurate results calibration curves were recorded and used for the determination of the equilibrium concentration of the alcohol c_A and each acceptor function of the peptide c_B. This approach demands that the extinction coefficient and hence the conformation in the peptide backbone does not change due to the addition of the alcohol which is fulfilled at low alcohol concentrations.

Assuming that the basicity of a polyfunctional acceptor is the combination of the separate K_a, K_b and $K_{c,}$ or from the decrease of the OH stretching signal according to eq. 16. This K will be named K_{OH}.

If the concentration ratio c_{A0}/c_{B0} gets larger than 1 the chances to form higher associates and species of the sort $B^a(A)B^b(A)B^c$, $B^a(A)B^bB^c(A)$, $B^aB^b(A)B^c(A)$ and $B^a(A)B^b(A)B^c(A)$ should be formed. From the statistical point of view the population of 1:1 associated species will be always low for the peptides since we still work in molecular concentrations lower than 0.02 mol/l. However, it should be proved experimentally by varying the peptide concentration with constant c_{A0}. Fig. 57 shows K as a function of c_{A0}/c_{B0}.

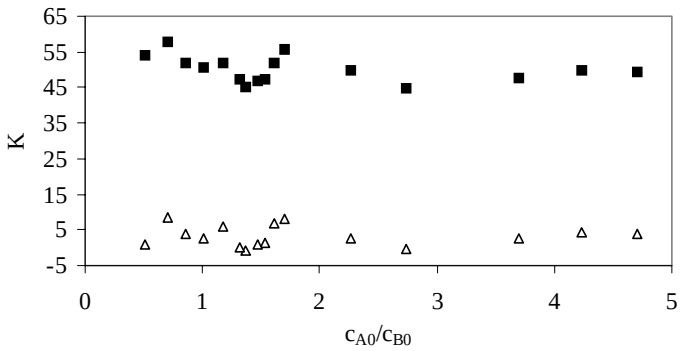


Fig. 57: Equilibrium constant K vs. c_{A0}/c_{B0} for the proton acceptor functions in Ac-Phe-OMe dissolved in methylene chloride at 25 °C (square: K_{Ac}, triangle: K_{Ester})

Two major conclusion can be drawn from the plot of the equilibrium constant vs. $c_{A0}/c_{B0.}$

The equilibrium constant is constant over a broad range of c_{A0}/c_{B0}. Hence, no 1:2 association occurs in excess concentration of HFiP. Surprisingly the ester function does not show any sign for an associate formation although it should be easily accessible for the proton donor. The equilibrium constant calculated for the association on this position is ca. zero.

Furthermore, the conformation effects due to the presence of HFiP is not relevant in the associate formation under the experimental conditions.

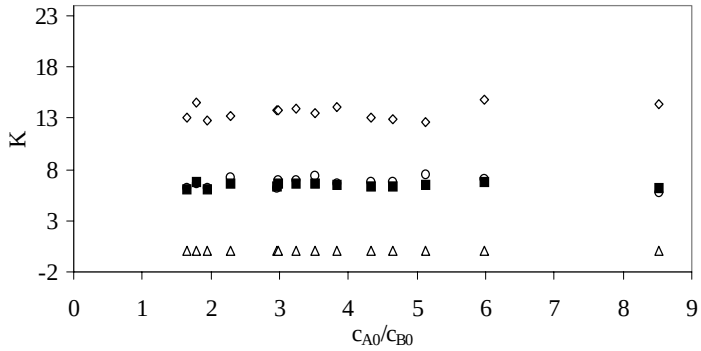


Fig. 58: Equilibrium constant K vs. c_{A0}/c_{B0} for the proton acceptor functions in Z-Ala-Val-OMe dissolved in methylene chloride at 25 °C (square: K_{all}, diamond: $K_{Peptide}$, circle: $K_{Urethane}$, triangle: K_{Ester})

The conclusions given above for the amino acid derivatives were confirmed by the plot of K vs.c_{A0}/c_{B0} of the association of HFiP with dipeptide and tripeptide derivatives as shown on the example of Z-Ala-Val-OMe in Fig. 58. The equilibrium constant does not change with increasing alcohol concentration and hence 1:2 association can be neglected in the discussion. However, association occurs on different acceptor positions e. g. the urethane and the peptide functions which exist in an equilibrium state. An interaction with the ester function could not be found. This effect allows us to use the ester carbonyl signal as an internal standard for the calculation. Attempts to adjust concentration differences by the correlation with the NH intensity failed because the extinction coefficient of the NH is obviously influenced by the hydrogen bond formation on the neighbour CO group. The standard deviation of the equilibrium constants given in tables 57 to 59 can be given as less than 10 %.

Based on the non-occurrence of higher associates the average equilibrium constant can be given in table 57 for the amino acid derivatives.

Table 57: Equilibrium constants of the association of acetyl amino acid methyl esters with alcohol molecules [in l/mol]

System	K_{OH}	K_{Ester}	$K_{Ac/Z}$
with HFiP			
Ac-Ala-OMe	86.4	0.0	49.0
Ac-Phe-OMe	84.2	0.0	46.1
Ac-Val-OMe	80.5	0.0	49.0
with Phenol			
Ac-Val-OMe	17.0	0.0	10.7

Comparing the equilibrium constants calculated from the decrease of the OH signal which gave K_{OH} and the K value calculated from the change of the intensity of the acceptor signal K_{Ac} large discrepancies become obvious. K_{OH} is almost twice as large as K_{Ac}. One explanation might be that besides association on these positions the electron lone pair of the nitrogen atom gets attacked. Also the π-electron system can act as acceptor function.

A further result concerns the peculiarities of the amino acid derivatives. Generally K_{Ac} and K_{OH} are similar in magnitude for all acetyl amino acid methyl esters. The small differences are rather caused by the changed basicity than by steric hindrance.

In the case of the dipeptides we observe an association on the urethane and the peptide function giving the equilibrium constants in table 58.

Table 58: Equilibrium constants of the dipeptide-alcohol interaction in methylene chloride [in l/mol]

System	K_{OH}	K_{Ester}	K_Z	$K_{Peptide}$	K_{all}
with HFiP					
Z-Ala-Ala-OMe	97.7	0.0	6.0	4.2	3.21
Z-Phe-Phe-OMe	81.2	0.0	4.9	6.8	3.87
Z-Val-Val-OMe	97.6	0.0	6.1	9.5	4.77
Z-Ala-Val-OMe	76.2	0.0	6.7	13.6	6.45
with Phenol					
Z-Val-Val-OMe	11.3	0.0	0.0	1.3	0.32

Comparing the K_{OH} data in table 58 with those in table 57 no significant changes were found. This confirms the conclusions drawn from Fig. 56 and 57 where no indication for a 1:2 hydrogen bond formation was observed. The acceptor function in the dipeptide compete for an associate formation with HFiP. Hence it can be understood why the equilibrium constants K_Z and $K_{Peptide}$ are significantly lower than K_{Ac}. The magnitude of the equilibrium constant should

be determined by two effects: the hydrogen bond basicity and the accessibility of the acceptor function. The relative shift of the acceptor signal due to the association can be taken as an indicator for the hydrogen bond basicity. For Z-Ala-Val-OMe the $\Delta\nu_{(COfree- COass)}$ values on the peptide function and the urethane function are 16.2 and 16.1 cm^{-1}, respectively, and confirm that the basicity of a urethane, the peptide and the acetyl amide function are comparable. Generally it would have been expected that the acceptor functions in the peptide backbone should be less accessible than functions on the C- and N- terminal end of the peptide. However, that does not mean that the equilibrium constants of an urethane function has to be larger than that of a peptide function. In fact this is only observed that $K_Z > K_{Peptide}$ in the case of Z-Ala-Ala-OMe. For the other dipeptide derivatives it is found that $K_Z < K_{Peptide}$.

Comparing the equilibrium constants K_{all} calculated according to eq. 31 of the dipeptide derivatives with K_{Ac} which represents the equilibrium constants for the association on the acceptor functions in the amino acid derivative $K_{all} < K_{Ac}$. Keeping in mind that the K_{OH} data were comparable this seems rather surprising but is can be explained by the different methods for the calculation. Hence in dipeptides we have to include the associate formation on all acceptor positions into the calculation whereas K_{OH} is exclusively calculated from the intensity changes of the OH signal.

Table 59: Equilibrium constants of the tripeptide-alcohol interaction in methylene chloride

System	K_{OH}	K_{Ester}	$K_{Ac/Z}$	$K_{Peptide}$	$K_{Peptide}$	K_{all}
with HFiP						
Z-Ala-Ala-Ala-OMe		0.0	8.8	5.9	13.0	10.7
Z-Phe-Phe-Phe-OMe	132.1	0.0	6.1	6.1	16.5	6.9
Z-Val-Val-Val-OMe	146.8	0.0	6.4	(-8.2)	21.2	3.6[a]
Z-Ala-Val-Val-OMe	216.3	0.0	4.9	7.7	12.1	5.9
with Phenol						
Z-Val-Val-Val-OMe	13.9	0.0	1.3	(-6.6)	7.3	0.4[a]

[a]The negative equilibrium constants in row 5 were not included in the calculation of K_{all}.

Enlarging the peptide chain by another amino acid residue the calculation becomes complicated because of the presence of conformational equilibria in the peptide bond. Thus, negative equilibrium constants for the association with Z-Val-Val-Val-OMe with phenol and HFiP are calculated from the intensity change of the signal at 1689.7 cm^{-1}. Obviously the association on one acceptor position in the molecule does cause a change in the peptide backbone which results in an increase of the extinction coefficient of the signal at 1689.7 cm^{-1}

and makes it impossible to receive reliable equilibrium constants for this molecule. Note that this happens independently on the proton donor used in the study only for Z-Val-Val-Val-OMe which is the only peptide with a branched side chain on C^{β}. The ^{1}H NMR titration indicated that the change in the conformation of the peptide backbone as a consequence of the hydrogen bond formation might be possible. The occurrence of negative equilibrium constants demonstrates the limitations of the approach to calculate the equilibrium constants from the decrease of the acceptor signal clearly.

Comparing K_{OH} with K_{all} the same effect is found as discussed for the peptide derivatives. Generally, $K_{OH}>K_{all}$ but with regards of the magnitude of the data the K_{OH} and K_{all} for the tripeptide derivatives are larger than for the dipeptide derivatives. In the group of the tripeptides the differences in K_{OH} and K_{all} are not significantly dependent on the kind of the amino acid residue in the side chain but rather by the acidity of the alcohol side in the investigation. Hence, K for the association with phenol are drastically lower than with HFiP.

The tendency of favouring the hydrogen bond formation is obvious: The preferred acceptor positions are the peptide functions. This occurs also when due to steric hindrance the acceptor positions are shielded. Thus, the conformational equilibrium will be changed. It would have been proposed that the urethane function is the best accessible and moderately basic group in the molecule and, thus, gives the larger K values. But this seems to be only the case for Z-Ala-Ala-OMe. For the ester function we have no hint that any hydrogen bonding takes place.

The question now arises what to expect when the peptide chain is further extended by one amino acid residue. In those cases the formation of the first loop of a helical structure can be realised under formation of a ten membered ring, C_{10}. The energetic stabilisation of the molecule is, however, not more effective than the C_7 rings. In the polar methylene chloride C_7 rings were not observed and for those reasons C_{10} rings should not occur either. As long as the attack of a proton donor on the preferred peptide acceptor functions is not hindered by the shielding of the apolar side chain residues the intramolecular hydrogen bonding will be inferior to the intermolecular interaction. Unfortunately, the enlargement of the peptide backbone does lower the solubility of the peptide in apolar solvents like methylene chloride drastically. Hence Ac-Ala-Ala-Ala-Ala-OMe is not sufficiently soluble to perform quantitative IR studies.

B.3.5 Correlation of the effective hydrogen bond basicity determined from chromatographic and infrared measurements

The equilibrium constants can be used to develop an hydrogen bond basicity parameter β_2^H according to eq. 15. The hydrogen bond acidity α_2^H for HFiP and phenol is 0.771 and 0.596, respectively. If calculated from the separate equilibrium constants the resulting hydrogen bond basicity β_2^H should be an additive property of the β_2^H values of each acceptor position. However, it should be noted here again that the hydrogen bond basicity β_2^H calculated from the equilibrium constant will not be of the same value as the effective hydrogen bond basicity $\sum\beta_2^H$ received form the HPLC experiment. This cannot be expected because the measurements conditions were completely different. In the HPLC experiment the compounds were fully hydrated whereas the peptides in the IR studies are surrounded by apolar solvent molecules. Consequently the conformation and the association properties will be different.

Table 60 summarises the hydrogen bond basicity β_2^H of the peptide derivatives calculated from K_{OH}. Furthermore, it gives a sum over the hydrogen bond basicities β_2^H of the separate acceptor functions.

Table 60: Hydrogen bond basicity scales β_2^H calculated from K_{OH} and the β_2^H of the separate acceptor positions

Substance	β_2^H(OH)	β_2^H(sum)
Ac-Ala-OMe	0.69	0.64
Z-Ala-Ala-OMe	0.69	1.37
Z-Ala-Ala-Ala-OMe	0.78	2.01
Ac-Phe-OMe	0.69	0.64
Z-Phe-Phe-OMe	0.68	1.40
Z-Phe-Phe-Phe-OMe	0.72	1.85
Ac-Val-OMe	0.69	0.64
Z-Val-Val-OMe	0.69	1.46
Z-Val-Val-Val-OMe	0.72	1.86
Z-Ala-Val-OMe	0.68	2.05
Z-Ala-Val-Val-OMe	0.75	1.98
with Phenol		
Ac-Val-OMe	0.77	0.71
Z-Val-Val-OMe	0.72	1.31
Z-Val-Val-Val-OMe	0.74	2.01

As it would have been expected the values in both columns are only similar for the protected amino acid derivatives since these act as monofunctional acceptor molecules. For di- and tripeptide derivatives the values received by adding β_2^H are ca. twice as large as β_2^H (OH). Discussing our results in the context of the β_2^H values the amino acid data seem reliable: For acetone, formamide, acetamide and tetramethylurea β_2^H values of 0.497, 0.662, 0.729 and 0.743, respectively, are given in the literature[253].

It was interesting for us to see if the β_2^H scale received from the infrared data gives similar values as the hydrogen bond basicity scales previously mentioned in this work. Fig. 59 compares different scales.

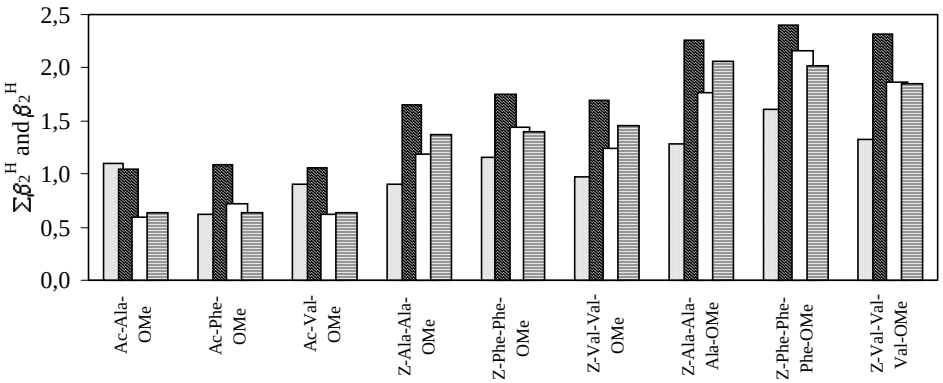


Fig. 59: Comparison of the hydrogen bond basicity scales received from different methods (grey: HPLC descriptors, black: calculated from the fragments of the UCL database, white: calculated from peptide fragments, dark grey: β_2^H (sum), respectively, see: Table A16)

Generally the tendency of a systematic increase of the hydrogen bond basicity of peptide derivatives is represented in all scales. Naturally, the experimental values will be determined by specific interactions with the solvent molecules or in the case of the HPLC scale with the stationary phase as well. Hence these experimental conditions will reflect on the correlation between $\sum\beta_2^H$ based on the CHI measurements and the β_2^H scales of the IR measurements. The regression coefficient is low and mainly for tripeptide derivatives the $\sum\beta_2^H$ values are smaller than β_2^H (Table 61). The experimental specifics are averaged when comparing the IR scale with fragmental scales.

Table 61: Linear correlation of the β_2^H with $\sum\beta_2^H$ and fragmental scales (y=m*x)

		$\sum\beta_2^H$ based on	
	HPLC measurements	UCL fragments	Amino acid fragments
m	0.7636	1.215	0.9113
r^2	0.1215	0.8617	0.8671

Comparing the regression coefficients the fragmental scales show an impressing correlation with the β_2^H scale. The first fragmental scale is based on 3 or 4 atomic fragments received from the analysis of the descriptors of the UCL database. As already described these hydrogen bond basicities are generally larger than the HPLC descriptors. Contrary the slope of the regression line in the correlation of $\sum\beta_2^H$ (HPLC) with β_2^H (IR) is smaller than one .

As a conclusion two different experimental methods e. g. HPLC gradient measurements and infrared measurements can be used for the determination of hydrogen bond basicity data. Both experiments will gives comparable but not identical values if it is possible to determine the equilibrium constants for each acceptor position individually. Because of its definition the equilibrium constant K_{OH} measured from the decrease of the OH signal is not useful for polyfunctional acceptor compounds. In any case it must be realised that only 1:1 associates occur.

C SUMMARY

For the drug development amino acid and oligopeptides represent an important class of compounds not just because they are applied as drugs themselves but also as substituents to create more body-like drugs and make them more tolerable to the human body. For the passive transport of these compounds through human tissues the lipophilicity/hydrophobicity of a drug represented by logP is essential but particularly the hydrogen bond properties need to be considered.

Initiated by these fundamental facts we have started a systematic investigation of structure-property relationship (QSPR) of amino acid, dipeptide and tripeptide derivatives which will allow the quantitative characterisation of the lipophilic/hydrophobic property of a peptide in general and the hydrogen bond acidity and basicity in particular. A quantitative structure-property relationship between the logarithm solute related property of a partitioning system and the structure of a solute is given by the solvation equation (eq. 4)[36]. Herein the five descriptors characterise the properties of a molecule which are essential in the solvent-solute interaction and are defined as excess molar refraction R_2, characteristic McGowan volume V_x, dipolarity/polarisability π_2^H, effective hydrogen bonding acidity $\sum\alpha_2^H$ and basicity $\sum\beta_2^H$. The six coefficients describe the contributions of the partitioning system to these particular interactions. Hence this equation not only provides a working tool for the prediction of logP data but also contains terms describing the hydrogen bond behaviour. With the help of this equation biological partitioning systems such as the water-skin partitioning and the blood-brain barrier were characterised by their coefficients thus allowing the prediction of $logP_{skin}$ and logBB for any compound with known descriptors.

Hence it was the intention of this work to develop descriptors for a number of oligopeptide and amino acid derivatives focusing on the hydrogen bond properties. Furthermore, these data should be used to create fragmental descriptors for amino acids and protecting functions which allow to approximate descriptors for any amino acid combination and oligopeptide sequence. The potential of this new fragmental descriptor system was tested on the hand of dipeptide and ureido sugar derivatives. In a second part conformational and associative effects were discussed on the hand of spectroscopic investigations.

The development of solute descriptors is usually based on the correlation of the logarithm of the partitioning coefficients of a compound measured in different partitioning systems by the traditional shake-flask method. Recently, Valko et al.[70] have introduced a new chromatographic parameter which characterises the hydrophobicity of a compound and can be

used as an alternative to $logP_{Oct}$. This so called chromatographic hydrophobicity index (CHI) can be received with a high precision from HPLC linear gradient retention times within 30 minutes and is, therefore, an attractive step towards high throughput methods.

Our strategy was the following: With the help of a test set containing 30 structurally unrelated compounds with known descriptors we have recorded the CHI values for 20 chromatographic partitioning systems. Among those there was also an immobilised artificial membrane (IAM) column which was believed to have the potential to simulate membrane properties. The regression analysis of the CHI values gave the six coefficients r, s, a, b, ν and c. As the comparison of the standardised coefficients with biological membrane has shown none of them represents the properties of the biological system. From the 20 chromatographic systems the five systems with most different coefficients have been chosen for the descriptor determinations. These are two classical reversed-phase ODS columns (In and Pro), a permethylated β-cylcodextrine phase (CD), a polar cyanopropyl phase (CN) and an immobilised artificial membrane column (IAM).

During the HPLC experiment the concentration of the organic modifier is increased linearly from 0 to 100 % within 3.5 minutes. As Hearn et al.[66] have shown in their studies the choice of the gradient time may influence the type of interaction of peptides and proteins with the active surface area. Since peptide analogue compounds were not among the 30 test compounds we investigated with great care the effect of temperature, gradient time and solvent polarity on the retention behaviour of the oligopeptide derivatives. As a conclusion no unexpected influence of the temperature and the gradient time on the retention time and bandwidth was observed which would indicate a hindered molecular flexibility of the peptide backbone. When using alcohols as organic modifiers instead of acetonitrile the retention behaviour is significantly different. In particular this is the case for 1.1.1-trifluoroethanol which is often recommended as excellent HPLC co-solvent due to its UV properties. However, we found that the solvation properties of TFE may effect the partitioning behaviour of peptide derivatives and will result in larger standard deviations in the calculation of the descriptors from CHI values. This peculiarity for the investigation of peptide derivatives can be explained on the hand of the infrared spectra: Small amount of fluorinated alcohols change the secondary structure of the peptide backbone.

The analysis of the CHI values regarding the effects of the peptide enlargement and the changes in the sequence have shown that the hydrophobicity represented by CHI is not an additive property when using the amino acid Gly as reference residue for the calculation of

amino acid contributions for amino acid, dipeptide and tripeptide derivatives consisting of the same amino acid residue. Furthermore, it was found that the CHI values are different if tripeptide derivatives with a variation in the amino acid sequence are compared.

The CHI values of the tripeptide derivatives have been used to determine amino acid contributions which allow a comparison with numerous amino acid hydrophobicity scales discussed in the literature. As a major conclusion it was found that our CHI scales correlate well with other scales either received from partitioning measurements of amino acid and amino acid derivatives or from HPLC measurements of synthetic octapeptides with systematic variation of the amino acid.

From the CHI values measured on the chromatographic columns In, Pro, CD, CN and IAM the descriptors for ca. 80 dipeptide and tripeptide derivatives were determined and discussed regarding the following aspects: peptide enlargement, amino acid substitution in the tripeptide derivatives Z-Ala-Xaa-Val-OMe and chirality of the amino acid.

In general, it was found that all substances with apolar side groups exhibit almost no effective hydrogen bond acidity. In part, this might be caused because we could not realise a wide variety of the coefficient a and therefore $\sum\alpha_2^H$ can be only given with a standard deviation of ±0.05, whereas these values for $\sum\beta_2^H$ and π_2^H are lower with ±0.01 and ±0.02, respectively. However, Abraham et al.[98] have shown using a different approach that $\sum\alpha_2^H$ for protected tripeptide derivatives will be expected as close to zero.

For Ac-Phe-OMe, Z-Phe-OMe, Z-Phe-Phe-OMe and Z-Phe-Phe-Phe-OMe the effective hydrogen bond basicity $\sum\beta_2^H$ is 0.62, 0.83, 1.16 and 1.61, respectively.

Regarding the amino acid substitution in tripeptide derivatives, Z-Ala-Xaa-Val-OMe, it was found that the effective hydrogen bond basicity and the dipolarity/polarisability of the compounds decreases with decreasing hydrophobicity. The opposite tendency was observed for the effective hydrogen bond acidity $\sum\alpha_2^H$. Exceptions from this general rule were the tripeptides with Xaa equals to Tyr and Ser. Here the polar OH functions increase the hydrogen bond acidity and compensate the basic properties of the compounds. Protection of the polar side functions in tripeptides by benzyl functions increases the hydrophobicity of the compound and its $\sum\beta_2^H$ values. The effect of diastereoisomerism on $\sum\beta_2^H$ is small but surprisingly persistent in all tripeptide derivatives: the effective hydrogen bond basicity was always lower for the LDL diastereoisomers than for the LLL and the DLL forms where the differences could be almost neglected. Infrared spectroscopically the equilibrium constant and association enthalpy for the intramolecular hydrogen bonding of the diastereoisomers of Z-

Ala-Phe-Val-OMe and Z-Ala-Leu-Val-OMe dissolved in carbon tetrachloride can be determined. Thus, K was 0.65, 0.85 and 0.77, for the LLL, LDL and DLL diastereoisomers, respectively. As a conclusion the hydrogen bond behaviour is different in diastereoisomers due to the accessibility of the acceptor functions.

Attempts by Platts to predict solute descriptors from fragments consisting of up to 4 atoms deduced from common organic molecules other than peptides did fail. We created fragmental solute descriptors for amino acids based on the analysis of the descriptors of 61 tripeptide derivatives. We could show that our fragmental descriptors can predict $\sum\beta_2^H$ (r^2=0.9113) and π_2^H (r^2=0.930) with high accuracy if the dipeptides consist exceptionally of apolar amino acid residues. The fragmental descriptors can be also applied for the prediction of descriptors for ureido sugar derivatives. Generally the prediction of $\sum\beta_2^H$ and π_2^H give satisfying agreements with the experimental data. Only for ureido sugar derivatives with -Gly-OEt, -D-Val-OEt, -Phe-Gly-OEt and -Ala-Phe-OBzl side chains the deviation was larger than 10% for $\sum\beta_2^H$. Infrared spectroscopic measurements could prove that these derivatives form strong intramolecular hydrogen bonds which might explain the deviation.

Although the solute descriptors can describe overall properties of a molecule in a complex manner specific interactions taking place in functions and conformational peculiarities in the molecule cannot be explained on the hand of the descriptors: Hence in the second part of the report we have focused our efforts on the spectroscopic investigation of the conformational and associative behaviour of amino acid, dipeptide and tripeptide derivatives. The effect of solvent polarity was of special interest. Among other solvents with either proton donor or acceptor or apolar properties such as D_2O, HFiP, DMSO-d6, CH_2Cl_2 and CCl_4, we have studied the behaviour of the peptide on the flexibility of a phospholipid DHPC which was chosen to simulate biological membranes and IAM column interactions.

Concluding from the position of the amide I signals of the amino acid derivatives dissolved in D_2O the bands were found red-shifted with increasing $\sum\beta_2^H$. This tendency was not found for the ester carbonyl signal. Obviously steric effects due to the bulkiness of the amino acid residues should be considered here as well. It should be noted that the amide I signal of the hydrated N-methyl acetamide is observed at 1623 cm^{-1} and, thus, only slightly shifted to lower wavenumbers compared to the amino acid derivative signals. Hence if the amide function can form hydrogen bonds of ideal symmetry regarding the distance and the bond angle the amide I band in peptide derivatives should be expected at similar wavenumbers. NMR studies have shown that the occurrence of *cis* isomers in the amide/peptide bond does not exceed an

amount of 2 % and can be, therefore, neglected in the further assignment of the amide I signals. Acetyl protected amino acid esters form C_5 intramolecular hydrogen bond associates in methylene chloride.

In di- and tripeptide derivatives a distribution of different conformers must be considered; the distribution function of the conformers will be strongly determined by the polar properties of the solvent and the capability of the peptide molecules to respond to the solvent interaction with changes in the peptide backbone. Hence the bulkiness of the amino acid side chains determine the arrangement. Based on this consideration we attempted an assignment of the signals. Intramolecular hydrogen bonds forming C_7 or C_{10} associates are distinguished in some rare cases in carbon tetrachloride solutions. VCD studies of the amide I region in this solvent have shown that for diastereoisomers of Z-Ala-Phe-Val-OMe and Z-Ala-Leu-Val-OMe an *trans*-C_7-C_7 arrangement would be energetically favoured and give an intense signal at ca. 1650 cm^{-1}. Furthermore, from *ab initio* MFP calculation the *trans* C_5-C_7 associates and the *trans* extended β- sheets were proposed as favoured arrangements in a model dipeptide with LL diastereoisomerism. In DL and LD diastereoisomers the occurrence of *trans* type II- β turns (C_{10} associates) becomes relevant. The equilibrium constant for the intramolecular association was 0.65, 0.85 and 0.77, for the LLL, LDL and DLL diastereoisomers of Z-Ala-Phe-Val-OMe, respectively, at 25 °C. The strength of the intramolecular hydrogen bonding affects the diffusion behaviour of the peptides: for the LDL diastereoisomer (0.17 mol/l*min) the diffusion rate was higher than for the LLL form (0.096 mol/l*min). In toluene the diffusion rate accelerate by 50 to 100 times.

Introducing tripeptide derivatives into a membrane like environment simulated by DHPC a strong dependency on the sequence of the tripeptide derivatives was found. In the hydrated 1:1 mixture of DHPC and Z-Ala-Phe-Val-OMe or Z-Ala-Val-Phe-OMe the flexibility of the membrane was drastically reduced and the transition from the gel P_β to the lamellar L_α could not be detected in the infrared spectra. Contrary, the addition of Z-Phe-Ala-Val-OMe increases the chain melting temperature by ca. 20 K. This effect might be induced by a modification of the crystalline structure of the tripeptide derivative.

The hydrogen bond basicity β_2^H can be also quantified based on the measurement of the equilibrium constant between a proton donor and proton acceptor molecule which form exclusively 1:1 associates under defined conditions (T=25 °C, CCl_4). Analysing $\log K_{1:1}$ data for associates measured in methylene chloride it was found that there exists also a linear correlation within a class of proton acceptors carrying X=O, C-Y-C functions (X=C, S, P, As,

Y=O, S). Another difficulty occurs when determining equilibrium constants in peptide derivatives because of the occurrence of more than one acceptor function in the molecule. Additionally peptide molecules contain NH functions as well which may act as proton donor in intramolecular hydrogen bonds. Due to the co-operative effect the acidity of these protons changes as a consequence of the intermolecular association. Hence it was necessary to develop an approach for the calculation of equilibrium constants in multifunctional proton donor/acceptor systems. As a model we have studied the hydrogen bond behaviour of propan-1.2-diol, propan-1.3-diol, butan-1.4-diol, pentan-1.5-diol, hexan-1.6-diol, cylcohexan-1.4-diol and cylcohexan-1.2-diol (*cis* and *trans*) with pyridine derivatives. With increasing number of CH_2 groups separating the OH function the intramolecular hydrogen bond acidity of the OH functions varies. Hence to receive the equilibrium constant of cyclohexan-1.4-diol, pentan-1.5-diol and hexan-1.6-diol with pyridine both hydroxyl groups can be treated as acting individually. With increasing amine concentration 1:2 association becomes more relevant; the equilibrium constant $K_{1:2}$ can be calculated based on $K_{1:1}$ received from measurements at low amine concentration (lower than 0.1 mol/l). In diols with the OH functions in 1.2, 1.3 or 1.4 position intramolecular association occurs which affects the acidity of the OH functions. In the infrared spectra of propan-1.2-diol the OH functions of different hydrogen bond acidity give separate signals at 3621.5 and 3533.7 cm^{-1}, for the free and the intramolecular associated OH species, respectively. Offering another acceptor compound to the solution of such diols both OH species, the free and the intramolecular OH associated, will form a hydrogen bond with the pyridine molecule.

We have developed an approach to describe the 1:1 and 1:2 associate formation statistically. A major conclusion from the statistics was that the number of 1:1 associates should not exceed an amount of half of the molecules in the solution. Choosing the amine concentration so that the intensity of the OH band after the addition of the acceptor molecules is just half of its initial concentration allows the calculation of equilibrium constants. Comparison with the statistical values shows that the experimental values are generally lower which is reasoned by electronic and steric effects.

The equilibrium constants of the amino acid and peptide derivatives with HFiP and phenol were calculated from the decrease of the potential acceptor signals which are for Z-Val-Val-Val-OMe located at 1740, 1723, 1688 and 1671.8 cm^{-1} for the ester carbonyl, the urethane and the two peptide signals. NMR titration have shown that the association of HFiP at a peptide acceptor function results in a high field shift of the NH signal of the urethane function and

hence causes a change in the conformation of the peptide backbone at alcohol concentrations above 0.01 mol/l which was avoided in the determination of the equilibrium constants. Furthermore, it was found that the ester carbonyl function was never attacked by HFiP or phenol. Changing the concentration ratio c_{A0}/c_{B0} by varying the peptide concentration c_{B0} in the range between 2 and 10 did not change the equilibrium constant for the separate acceptor positions within a standard deviation of 10 %. Hence 1:2 associates can be excluded. Comparing the acceptor functions according to their tendency to form hydrogen bonds with HFiP the preferred functions are the peptide groups although it would have been expected the these functions are well shielded against the alcohol attack. The effect of the bulkiness of the amino acid residue becomes evident in Z-Val-Val-Val-OMe. For this peptide a negative equilibrium constant was calculated for one peptide function. Obviously the branching on C^{β} demands a change in the backbone conformation for an effective alcohol attack. The hydrogen bond basicity β_2^H calculated from logK at the separate acceptor functions can be given as 0.64, 1.40 and 1.85, for Ac-Phe-OMe, Z-Phe-Phe-OMe and Z-Phe-Phe-Phe-OMe, respectively. The β_2^H data are of similar magnitude as the $\sum\beta_2^H$ values received from other methods. Best correlation was found with the $\sum\beta_2^H$ scale calculated based on the amino acid fragments. Differences between both experimental scales are not surprising because the conditions were not comparable: In the HPLC experiment the proton donor were present in a large excess whereas in the infrared experiment the acceptor groups competed for the donor molecules in order to form 1:1 associates. Fragmental scales do not cover the peculiarities of these interactions.

D REFERENCES

1. IUPAC-IUB Commission on Biochem. Nomenclature, *Biochemistry,* **9** (1970) 3471
2. A. N. Eberle, *Chimia* **45** (1991) 145
3. J. M. van Ree, *J. Controlled Release,* **29** (1994) 307
4. S. M. Edgington, *Biotechnology,* **9** (1991) 1327
5. E. M. Wright, *Ann. Rev. Respir. Dis.,* **127** (1983) S3
6. T. Terasaki, A. Tsuji, *J. Controlled Release,* **29** (1994) 163
7. I. Tamai, A. Tsuki, *Adv. Drug Delivery Reviews,* **19** (1996) 401
8. D. J. Begley, in: M. W. B. Bradbury (Ed.) *Handbook of Experimental Pharmacology, Physiology and Pharmacology of the Blood-Brain Barrier,* Springer, Berlin, (1992) pp. 151
9. B. V. Zlokovic, M. N. Lipovac, D. J. Begley, H. Davson, Lj. Rakic, *J. Neurochem.,* **49** (1987) 310
10. B. V. Zlokovic, S. Hyman, J. G. McComb, M. N. Lipovac, G. Tang, H. Davson, *Biochem. Biophys. Acta,* **1025** (1990) 191
11. B. V. Zlokovic, M. N. Lipovac, D. J. Begley, H. Davson, Lj. Rakic, *J. Neurochem.,* **51** (1988) 252
12. D. J. Begley, *J. Controlled Release,* **29** (1994) 293
13. R. N. Smith, C. Hansch, M. M. Ames, *J. Pharm. Sci.* **64** (1975) 599
14. A. Leo, C. Hansch, D. Elkins, *Chem. Rev.* **71** (1977) 525
15. S. J. Singer, G. L. Nicolson, *Science* **175** (1972) 720
16. P. S. Burton, R. A. Conradi, A. R. Hillgers, N. F. H. Ho, L. L. Maggiora, *J. Controlled Release* **19** (1992) 87
17. J. Comer, *Chem. in Britain* **30** (1994) 983
18. A. Avdeef, *J. Pharm. Sci*. **82** (1993) 1
19. C. Hansch, A. Leo, *Chem. Rev.* **91** (1993) 1281
20. R. Mannhold, H. Kubinyi, H. Timmerman, in: V. Pliska, B. Testa, H. van de Waterbeemd (Eds.); *Lipophilicity in Drug Action and Toxicology,* VCH Weinheim. New York, (1996), pp. 141
21. C. Hansch, T. Fujita, *J. Am. Chem. Soc.* **86** (1964) 1616
22. C. Hansch, P. P. Maloney, T. Fujita, R. M. Muir, *Nature (London),* **194** (1962) 178
23. T. Fujita, J. Iwasa, C. Hansch, *J. Am. Chem. Soc.,* **86** (1964) 5175
24. P. Broto, G. Moreau, C. Vandycke, *Eur. J. Med. Chem.* **19** (1984) 71
25. A. K. Ghose, G. M. Crippen, *J. Comput. Chem.* **7** (1986) 565
26. K. S. Roger, A. Cammarata, *Biochim. Biophys. Acta,* **193** (1969) 22
27. L. P. Hammett, *Physical Organic Chemistry,* McGraw-Hill New York (1940)
28. G. G. Nys, R. F. Rekker, *Chim. Theor.* **8** (1973) 521
29. D. J. Abraham, A. J. Leo, *Proteins, Struct. Funct, Gen. 2,* (1987) 130
30. W. J. Dunn III., S. Wold, *Acta Chem. Scand.,* **B32** (1978) 536
31. J. Janin, *Nature,* **277** (1979) 491
32. J. D. Cramer, III. *J. Am. Chem. Soc.,* **99** (1977) 5408
33. M. J. Kamlet, J. L. Abbroud, M. H. Abraham, R. W. Taft, *J. Org. Chem.,* **48** (1983) 877
34. M. J. Kamlet, J. L. Abbroud R. W. Taft, *J. Am. Chem. Soc.,* **99** (1977) 6027
35. A. Leo, *Chem. Rev.,* **93** (1993)1281
36. M. H. Abraham, *Chem. Soc. Revs.,* **22** (1993) 73
37. D. A. Paterson, R. A. Conradi, A. R. Hilgers, T. J. Vidmar, P. S. Burton, *Quant. Struct.-Act. Relat.,* **13** (1994) 4

38. M. H. Abraham, H. S. Chadha, G. S. Whiting, R. C. Mitchell, *J. Pharm. Sci.,* **83** (1994) 1085
39. M. H. Abraham, K. Takacs-Novak, R. C. Mitchell, *J. Pharm. Sci.,* **86** (1997) 310
40. M. H. Abraham, H. S. Chadha, R. C. Mitchell, *J. Pharm. Pharmacol.,* **47** (1995) 8
41. M. H. Abraham, S. H. Chadha, J. P. Dixon, C. Rafols, C. Treiner, *J. Chem. Soc. Perkin Trans. 2,* (1995) 887
42. H. J. Schaffer, L. Beauchamp, P. de Miranda, G. B. Elion, D. J. Bauer, P. Collins, *Nature***, 272** (1978) 583
43. A. S. Gilbert, J. Moss, P. L. Francis, M. J. Ashton, D. S. Ashton, *Chromatographia,* **42** (1996) 305
44. A. Kristl, S. Pecar, *J. Med. Chem.,* **32** (1997) 3
45. K. Szczepaniak, M. Szczesniak, W. Szajda, W. B. Person, J. Leszczynski, *Can. J. Chem.,* **69** (1991) 1705
46. K. Szczepaniak, M. Szczesniak, W. B. Person, *Chem. Phys. Lett*., **153** (1988) 39
47. K. Szczepaniak, M. Szczesniak, *J. Mol. Struct.,* **156** (1987) 29
48. P. R. LeBurton, X. Yang, S. Urano, S. Fetzer, M. Yu, N. J. Leonard, S. Kumar, *J. Am. Chem. Soc.* **112** (1990) 2138
49. A. Sygula, A. Buda, *Theochem.,* **9** (1983) 267
50. L. Gorb, J. Leszczynski, *Int. J. Quant. Chem.,* **64** (1997) 759
51. J. Leszczynski, *Chem. Phys. Lett.,* **174** (1990) 347
52. L. B. Clark, I. Tinoco, *J. Am. Chem. Soc.,* **87** (1965) 11
53. B. Albinson, B. Norden, *J. Am. Chem. Soc.,* **115** (1993) 223
54. H. Gao, A. K. Mitra, *Magn. Reson. Chem.,* **37** (1999) 687
55. R. Collander, *Acta Physiol. Scand*, **13** (1947) 363
56. A. Leo in: R. G. Gould (Ed.) *Biological Correlations-The Hansch Approach* (*Advances in Chemistry* Ser. *114)* American Chem. Soc., Washington. DC. (1972) pp. 51
57. M. S. Mirrees, S. J. Moutin, C. T. Murphy, P. J. Taylor, *J. Med. Chem*., **19** (1979) 619
58. S. H. Unger, J. R. Cook, J. S. Hollenberg, *J. Pharm. Sci.,* **67** (19778) 1364
59. Th. Braumann, G. Weber, L. H. Grimme, *J. Chromatogr.,* **261** (1983) 329
60. W. J. Lambert, *J. Chromatogr.,* **656** (1993) 469
61. M. H. Abraham, H. S. Chadha, G. S. Whiting, R. C. Mitchell, *J. Pharm. Sci.,* **83** (1994) 1085
62. K. G. Miller, C. F. Poole, *J. High Resolut. Chromatogr.,* **17** (1994) 125
63. M. H. Abraham, M. Roses, *J. Phys. Org. Chem.,* **7** (1994) 672
64. K. Valko, P. Slegel, *J. Chromatogr.,* **631** (1993) 49
65. A. W. Purcell, M. I. Aguilar, M. T. W. Hearn, *J. Chromatogr.,* **593** (1992) 103
66. A. W. Purcell, M. I. Aguilar, M. T. W. Hearn, *Analyt. Chem.,* **65** (1993) 3038
67. R. M. Smith, *J. Chromatogr.,* **236** (1982) 313
68. R. M. Smith, *J. Chromatogr.,* **291** (1984) 372
69. J. J. Pesek, E. Williamsen in: R. M. Smith (Ed.) *Retention and Selectivity in Liquid Chromatography*, *Journal of Chromatography Library*, Vol. 57. Elsevier Science B.V. (1995), pp. 3
70. K. Valko, C. Bevans, D. Reynolds, *Analyt. Chem.,* **69** (1997) 2022
71. P. Kuronen in: R. M. Smith (Ed.) *Retention and Selectivity in Liquid Chromatography, Journal of Chromatography Library.* Vol. 57. Elsevier Science B.V. (1995), pp. 20
72. A. Kaibara, M. Hirose, T. Nakagawa, *Chromatographia*, **29** (1990) 59
73. C. M. Du, K. Valko, C. Bevan, D. Reynolds, M. H. Abraham, *Analyt. Chem.,* **70** (1998) 4228
74. M. H. Abraham, UCL database, (1997)

75. M. H. Abraham, G. S. Whiting, R. M. Doherty, W. J. Shuley, *J. Chem. Soc. Perkin. Trans. 2,* (1990) 1451
76. M. H. Abraham, J. C. McGowan, *Chromatographia,* **23** (1987) 243
77. Cs. Horvath, W. R. Melander, I. Molnar, *J. Chromatogr.,* **125** (1976) 129
78. W. R. Melander, D. Corradini, Cs. Horvath, *J. Chromatogr.,* **317** (1984) 67
79. W. R. Melander, Cs. Horvath, *Arch. Biochem. Biophys.,* **183** (1977) 393
80. F. Beigi, P. Lundahl, *J. Chromatogr.,* **852** (1999) 313
81. C. Pidgeon. U. V. Venkatarum, *Anal. Biochem.,* **176** (1989) 36
82. A. W. Purcell, M. I. Aguilar, M. T. W. Hearn, *J. Chromatogr.,* **593** (1992) 103
83. A. W. Purcell, G. L. Zhao, M. I. Aguilar, M. T. W. Hearn, *J. Chromatogr*., **852** (1999) 43
84. M. I. Aguilar, S. Mougos, J. Boublik, J. Rivier, M. T. W. Hearn, *J. Chromatogr.*, **646** (1993) 53
85. A. W. Purcell, M. I. Aguilar, M. T. W. Hearn, *J. Chromatogr*., **711** (1995) 71
86. C. T. Mant, R. S. Hodges, *J. Liquid Chromatogr*., **12** (1989) 139
87. L. R. Snyder, in: Cs. Horvath (Ed.) *HPLC-Advances and Perspectives,* Academic Press, New York, (1980) pp. 207
88. D. Guo, C. T. Mant, A. K. Taneja, J. M. R. Parker, R. S. Hodges, *J. Chromatogr.,* **359** (1986) 499
89. J. M. R. Parker, D. Guo, R. S. Hodges, *Biochemistry,* **25** (1986) 5425
90. G. D. Rose, A. R. Geselowitz, G. J. Lesser, R. H. Lee, M. H. Zehfus, *Science,* **229** (1985) 834
91. W. R. Krigbaum, A. Komoriya, *Biochim. Biophys. Acta,* **576** (1979) 204
92. J. L. Cornette, K. B. Cease, H. Margalit, J. L. Spouge, J. A. Berzofsky, C. Delisi, *J. Mol. Biol.,* **195** (1987) 659
93. J. J. Yang, M. Pitkeathly, S. E. Radford, *Biochemistry,* **33** (1994) 7345
94. J. Zukowski, M. Pawlowska, M. Nagatkina, D. W. Armstron, *J. Chromatogr*., **629** (1993) 169
95. D. Guo, C. T. Mant, A. K. Taneja, R. S. Hodges, *J. Chromatogr.,* **359** (1986) 519
96. C. T. Mant, T. W. L. Burke, J. A. Black, R. S. Hodges, *J. Chromatogr.,* **458** (1988) 193
97. C. M. Du, pivate communication
98. M. H. Abraham, F. Martins, R. C. Mitchell, C. J. Salter, *J. Pharm. Sci.,* **88** (1999) 241
99. J. A. Platts, D. Butina, M. H. Abraham, A. Hersey, *J. Chem. Inf. Comput. Sci*., (1999) in press
100. G. Klopman, J.-Y. Li, S. Wang, M. Dimayuga, *J. Chem. Inf. Comput. Sci.,* **34** (1994) 752
101. C. Tanford, *The Hydrophobicity Effect,* 2nd ed. Wiley -Interscience, John Wiley &. Sons, New York, (1980)
102. A. Ben-Naim, *Hydrophobic Interactions,* Plenum Press, New York, (1980)
103. Y. Nakai, A. Kidera, M. Kanehisa, *Prot. Engineering,* **2** (1988) 93
104. J. M. R. Parker, R. S. Hodges, in: C. T. Mant, R. S. Hodges (Eds.) *High performance Liquid Chromatography, of Peptides and proteins: Separation Analysis and Conformation,* CRS Press. Boca Rato, (1991)
105. Y. Nozaki, C. Tanford, *J. Biol. Chem.,* **246** (1971) 2211
106. J. L. Fauchere, V. Pliska, *Eur. J. Med. Chem.,* **18** (1983) 369
107. G. D. Rose, L. M. Mirasch, J. A. Smith, *Adv. Protein Chem.,* **37** (1985) 1
108. M. C. J. Wilce, M.-I. Aguilar, M. T. W. Hearn, *Analyt. Chem.,* **67** (1995) 1210
109. M. C. J. Wilce, M. I. Aguilar, M. T. W. Hearn, *J. Chromatogr.,* **548** (1991) 117
110. J. L. Meek, Z. L. Rossetti, *J. Chromatogr.,* **211** (1981) 15

111. R. A. Houghten, S. T. DeGraw, *J. Chromatogr*., **386** (1987) 223
112. M. Zimmerman, N. Eliezer, R. Simha, *J. Theor. Biol*., **21** (1968) 170
113. H. B. Bull, K. Breese, *Arch. Biochem. Biophys*., **161** (1974) 665
114. C. Chlothia, *J. Mol. Biol.*, **105** (1976) 1
115. H. R. Guy, *Biophys. J*., **47** (1985) 61
116. G. von Heijne, C. Blomberg, *Eur. J. Biochem.*, **97** (1979) 175
117. R. Wolfenden, L. Anderson, P. M. Cullius, C. C. B. Southgate, *Biochemistry,* **20** (1981) 849
118. D. Eisenberg, R. M. Weiss, T. C. Terwilliger, W. Wilcox, *Faraday Symp. Chem. Soc.*, **17** (1982) 109
119. J. L. Meek, *Proc. Natl. Acad. Sci., U.S.A.*, **77** (1980) 1632
120. M. C. J. Wilce, M. I. Aguilar, M. T. W. Hearn, *Analyt. Chem.*, **67** (1995) 1210
121. S. P. Gupta, *Chem. Revs.*, **89** (1989) 1765
122. H. Schinke, Ph.D. thesis, Martin-Luther-Univeristy Halle-Wittenberg (1995)
123. R. R. Herr, H. K. Jahnke, A. D. Argoude-Lis, *J. Am. Chem. Soc.,* **89** (1967) 4808
124. J. J. Vavra., C. DeBoer, A. Dietz, L. J. Hanka, W. T. Sokolski, *Antibiot. Ann.*, (1959/1960) 230
125. E. J. Hessler, H. K. Jahnke, *J. Org. Chem.,* **32** (1970) 245
126. T. P. Johnston, G. S. McCaleb, J. M. Montgomerey, *J. Med. Chem.,* **18** (1975) 104
127. T. Anderson, M. McMenamin, P. S. Schein, *Cancer Res*., **35** (1975) 761
128. M. J. Kamlet, R. W. Taft, *J. Am. Chem. Soc.,* **98** (1976) 377
129. M. W. Taft, *J. Chem. Soc. Perkin Trans. 2,* (1979) 1723
130. C. Reichardt, *Angew. Chem. Int. Ed. Engl.*, **4** (1965) 29
131. E. Grunwald, S. Winstein, *J. Am. Chem. Soc*., **70** (1948) 846
132. V. Gutmann, *The Donor-Acceptor Approach to Molecular Interactions*, Plenum Press, New York, (1978) 34344
133. W. A. P. Luck in: P. L. Hyuskens, W. A. P. Luck, Th. Zeegers-Huyskens (Eds.), *Intermolecular Forces*; Springer Verlag, Berlin (1991) 41
134. M. H. Abraham, P. L. Grellier, D. V. Prior, J. J. Morris, P. T. Taylor, C. Laurence, M. Berthelot, *Tetrahedron Letters*, **30** (1989) 2571
135. M. H. Abraham, P. L. Grellier, D. V. Prior, P. P. Duce, J. J. Morris, P. J. Taylor, *J. Chem. Soc. Perkin Trans 2,* (1989) 699
136. M. H. Abraham, P. P. Duce, D. V. Prior, D. G. Barratt, J. J. Morris, P. J. Taylor, *J. Chem. Soc. Perkin Trans 2,* (1989) 1355
137. M. H. Abraham, P. L. Grellier, D. V. Prior, J. J. Morris, P. J. Taylor, *J. Chem. Soc. Perkin Trans 2,* (1990) 521
138. M. H. Abraham, P. L. Grellier, D. V. Prior, J.J. Morris, P. J. Taylor, R. M. Doherty, *J. Org. Chem*.**, 55** (1989) 2227
139. M. H. Abraham, P. L. Grellier, D. V. Prior, J. J. Morris, P. J. Taylor, C. Laurence, M. Berthelot, R. M. Doherty, M. J. Kamlet, J.-L. M. Abboud, K. Sraidi, F. Guiheneuf, *J. Am. Chem. Soc.,* **110** (1988) 8534
140. C. Laurence, M. Berthelot, M. Helbert, K. Sraidi, *J. Phys. Chem.,* **93** (1989) 3799
141. J-L. M. Abboud, K. Sraidi, M. H. Abraham, R. W. Taft, *J. Org. Chem.,* **55** (1990) 2230
142. J.A. Smith, W. L. Durax, D. A. Langs, G. T. DeTitta, J. W. Edmonds, D. C. Rohrer, C. M. Weeks, *J. Am. Chem. Soc.*, **97** (1975) 7242
143. L. M. Jackman, S. Sternhell, *Applications of Nuclear Magnetic Resonance Spectroscopy in Organic Chemistry*. Pergamond, Oxford, (1969)
144. E. L. Mehler, G. Eichele, *Biochemistry*, **23** (1984) 3887
145. C. Reichardt, *Solvents and Solvent Effects in Organic Chemistry*, VCH Weinheim, (1990) pp. 408

146. J. Murto, A. Kivinen, S. Kivimaa, R. Laakso, *Suomen Kemist.,* **40** (1967) 250
147. T. Miyazawa, T. Shimanouchi, S. I. Mizushima, *J. Chem. Phys.,* **24** (1956) 408
148. H. Susi, *Methods Enzymol.,* **26** (1972) 445
149. L. J. Bellamy, W. Brügel, *Ultrarot-Spektrum und chemische Konstitution,* Dr. Dittrich Steinkopff Verlag Darmstadt, (1966) p. 156
150. M. Jackson, P. I. Haris, D. Chapman, *Biochemistry,* **30** (1991) 9681
151. J. Reed, V. Kinzel, *Biochemistry,* **30** (1991) 4521
152. J. J. Yang, M. Pitkeathly, S. E. Radford, *Biochemistry,* **33** (1994) 7345
153. N. Hirota, K. Mizuni, Y. Goto, *Protein Sci.,* **6** (1997) 416
154. J. Wei, G. D. Fasman, *Biochemistry,* **34** (1995) 6408
155. Y. Wang, R. Purrelo, T. G. Spiro, *J. Am. Chem. Soc.,* **111** (1989) 8274
156. R. W. Williams, *Biopolymers,* **32** (1992) 829
157. T. Higashijama, M. Tasumi, T. Miyazawa, *Biopolymers,* **16** (1977) 1259
158. M. Branik, H. Kessler, *Chem. Ber.,* **108** (1975) 2176
159. J. W. Bats, H. Fuess, H. Kessler, R. Schuck, *Chem. Ber.,* **113** (1980) 520
160. A. C. Chermovitz, T. B. Freedman, L. A. Nafie, *Biopolymers,* **26** (1987) 1879
161. D. S. Dwyer, *Biopolymers,* **49** (1999) 635
162. M. Jackson, H. H. Mantsch, *Biochem. Biophys. Acta,* **1078** (1991) 231
163. M. Hollosi, K. E. Kover, S. Holly, L. Radics, G. D. Fasman, *Biopolymers,* **26** (1987) 1527
164. B. Imperiali, S. L. Fisher, R. A. Moats, T. J. Prins, *J. Am. Chem. Soc.,* **114** (1992) 3182
165. M. Plass, A. Kolbe, *J. Mol. Struct.,* **322** (1994) 241
166. M. Plass, Ph.D. Thesis Martin-Luther-University Halle-Wittenberg (1993)
167. F. R. Maxfield, S. J. Leach, E. R. Stimson, S. P. Powers , H. A. Scheraga, *Biopolymers,* **18** (1979) 2507
168. P. I. Haris, C. Chapman, *Biopolymers,* **37** (1995) 251
169. M. Jackson, P. I. Haris, D. Chapman, *Biochim. Biophys. Acta,* **998** (1989) 75
170. H. H. Mantsch, D. J. Moffatt, H. L. Casal, *J. Mol. Struct.,* **173** (1988) 285
171. W. K. Surewicz, H. H. Mantsch, *Biochem. Biophys. Acta,* **952** (1988) 115
172. D. M. Byler, H. Susi, *Biopolymers,* **25** (1986) 469
173. E. Goormaghtigh, R. Brasseur, P. Huart, J. M. Ruysschaert, *Biochemistry*, **26** (1987) 1789
174. L. Bragg, J. C. Kendrew, M. F. Perutz, *Proc. Royal Soc. London, Ser. A,* **203** (1950) 321
175. E. J. Miller, B. M. Ross, R. Ismail, K. Belhadj-Mostefa, R. Poet, *J. Mol. Biol.,* **204** (1988) 777
176. S. S. Zimmermann, M. S. Pottle, G. Nemethy, A. H. Schegara, *Macromolecule*s, **10** (1977) 1
177. A. Perczel, J. G. Angyan, M. Kaitar, W. Viviani, J.-l: Rivail, J.-F. Marcoccia, I. G. Csizmadia, *J. Am. Chem. Soc.,* **113** (1991) 6256
178. H. J. Dyson, M. Rance, R. A. Houghton, R. A.. Lerner, P. E. Wright, *J. Mol. Biol.,* **201** (1988) 161
179. P. E. Wright, H. J. Dyson, R. A. Lerner, *Biochemistry*, 27 (1988) 7167
180. Ch. Pulla Rao, P. Balaram, C. N. Rao, *Biopolymers,* **22** (1983) 2091
181. J. Parmentier, C. Samyn, M. Van Beylen, Th. Zeegers-Huyskens, *J. Chem. Soc. Perkin Trans 2,* (1991) 387
182. J. Parmentier, K. De Wael, Th. Zeegers-Huyskens, *J. Mol. Struct.,* **320** (1992) 67
183. H. H. Mantsch, A. Perczel, M. Hollosi, G. D. Fasman, *Biopolymers,* **33** (1993) 201
184. R. K. Konat, B. Mathä, J. Winkler, H. Kessler, *Liebigs Ann.,* (1995) 765

185. M. H. Rao, W. Yang, H. Joshua, J. M. Becker, F. Naider, *Int. J. Pept. Protein Res.,* **45** (1995) 418
186. G. Zanotti, M. Saviano, G. Saviano, T. Tancredi, F. Rossi, C. Pedone, E. Benedetti, *J. Peptide Res.,* **51** (1998) 460
187. E. Vass, M. Kurz, R. K. Konat, M. Hollosi, *Spectrochimica Acta A,* **54** (1998) 773
188. K. Tonan, S. Ikawa, *J. Am. Chem. Soc,* **1198** (1196) 6960
189. I. Z. Siemion, P. Stefanowicz, C. Czapoewiski, J. Ciarkowiski, *Polish J. Chem.,* **69** (1995) 902
190. M. Kawai, G. Fasman, *J. Am. Chem. Soc.,* **100** (1978) 3630
191. R. W. Woody in: K. Nakanishi, N. Berova, R. W. Woody, (Eds.) *Circular Dichroism: Principles and Applications,* VCH Weinheim, (1994), pp. 473
192. P. Lagant, G. Vergoten, G. Fleury, M.-H. Loucheux-Lefebvre, *Eur. J. Biochem.,* **139** (1984) 149
193. J. Bandekar, *Biochim. Biophys. Acta,* **1120** (1992) 123
194. M. Hollosi, Zs. Majer, A. Z. Ronai, K. Medzihradszky, S. Holly, *Biopolymers,* **34** (1994) 177
195. T. C. Cheam, S. Krimm, *J. Mol. Struct. (Theochem),* **188** (1989) 15
196. M. Hollosi, K. E. Köver, S. Holly, L. Radic, G. D. Fasman, *Biopolymers,* **26** (1987) 1555
197. W. C. Johnson Jr., I. Tinoco, Jr., *J. Am. Chem. Soc.,* **94** (1972) 4389
198. J. S. Balcerski, E. S. Pysh, G. M. Borona, C. Toniolo, *J. Am. Chem. Soc.,* **98** (1976) 3470
199. M. M. Kelly, E. S. Pysh, G. M. Bonora, C. Toniolo, *J. Am. Chem. Soc.,* **99** (1977) 3264
200. C. Toniolo, G. M. Bonora, *Can. J. Chem.,* **54** (1976) 70
201. L. Schäfer, S. Q. Newton, M. Cao, C. Peeters, C. Van Alsenoy, K. Wolniki, F. A. Momany, *J. Am. Chem. Soc,* **155** (1993) 272
202. M. H. Baron, C. De Lorze, C. Toniolo, G. D. Fasman, *Biopolymers,* **17** (1978) 2225
203. J. Kobajashi, G. Nagai, G. *Biopolymers,* **17** (1978) 2265
204. V. F. Bystrov, S. L. Portnova, A. Balashova, V. I. Tsetli, V. T. Ivanov, P. V. Kostezky, Yu. A. Ovchinnikov, *Tetrahedron Letters,* (1969) 5225
205. L. A. Bodack, T. B. Freedman, M. Plass, L. A. Nafie, in: M. Tasumi, K. Itoh (Eds.) *Proceedings of the 12th International Conference on FT Spectroscopy,*. submitted
206. T. B. Freedman, L. A. Nafie, T. A. Keiderling, *Biopolymers,* **37** (1995) 265
207. P. Bour, T. A. Keiderling, *J. Am. Chem. Soc.,* **115** (1993) 9602
208. B. A. Cornell, F. Separovic, *Biochim. Biophys. Acta,* **773** (1983) 189
209. D. Marsh, CRC *Handbook of Lipid Bilayers* CRC Press, (1990)
210. R. Koynova, M. Caffrey, *Biochim. Biophys. Acta,* **1376** (1998) 91
211. D. G. Cameron, H. L. Casal. H. H. Mantsch, Y. Boulanger, I. C. P. Smith, *Biophys. J.,* **35** (1981) 1
212. J. Boggs, G. Rangaraj, A. Watts, *Biochim. Biophys. Acta,* **981** (1989) 243
213. P. Laggner, K. Lohner, G. Degovics, K. Muller, P. Lagner, *Chem. Phys. Lipids,* **44** (1987) 31
214. J. Kim, J. Mattai, G. Shipley, *Biochemistry,* **26** (1987) 6592
215. M. Szogyi, T. Cserhati, B. Borgas, *Mol. Cryst. Liq. Cryst.,* **152** (1987) 267
216. R. E. Jacobs, S. H. White, *Biochemistry,* **25** (1986) 2605
217. R. E. Jacobs, S. H. White, *Biochemistry,* **26** (1987) 6127
218. M. Plass, I. Wawer, B. Pierkarska-Bartoszewicz, A. Temeriusz, *J. Phys. Org. Chem*., **10** (1997) 747
219. I. Wawer, B. Piekarska-Bartoszewicz, A. Temeriusz, *Carbohydr. Res.,* **267** (1995) 167

220. A. M. McManus, L. Otvos, Jr., R. Hoffmann, D. J. Craik, *Biochemistry*, **38** (1999) 705
221. A. H. Andreotti, D. Kahne, *J. Am. Chem. Soc.*, **115** (1993) 3352
222. R. Liang, A. H. Andreotti, D. Kahne, *J. Am. Chem. Soc.*, **117** (1995) 10395
223. E. Vass, E. Lang, J. Samu, Zs. Majer, M. Kajtar-Peredy, M. Mak, L. Radics, M. Hollosi, *J. Mol. Struct.*, **440** (1998) 59
224. M. Hollosi, A. Perczel, G. D. Fasman, *Biopolymers*, **29** (1990) 1549
225. S. Horvat, A. Jakas, E. Vass, J. Samu, M. Hollosi, *J. Chem. Soc. Perkin Trans.* 2, (1997) 1523
226. R. Anulewicz, I. Wawer, B. Piekarska-Bartoszewicz, A. Temeriusz, *Carbohydr. Chem*., **16** (1997) 739
227. M. Plass, I. Wawer, B. Pierkarska-Bartoszewicz, A. Temeriusz, *J. Phys. Org. Chem*., **10** (1997) 747
228. A. Wawer, I. Wawer, B. Piekarska-Bartoszewicz, A. Temeriusz, *Spectroscopy Lett*., **29** (1996) 1079
229. R. Anulewicz, I. Wawer, B. Piekarska-Bartoszewicz, A. Temeriusz, *Carbohydr. Chem*., **16** (1997) 739
230. M. Plass, M. Weychert, I. Wawer, B. Piekarska-Bartoszewicz, A. Temeriusz, *Carbohydr. Chem.*, (2000) in press
231. M. Weychert, J. Klimkiewicz, I. Wawer, B. Piekarska-Bartoszewicz, A. Temeriusz, *Magn. Res. Chem.,* **36** (1998) 727
232. A. Kolbe, H. Pracejus, *Ber. Bunsengesell. Phys. Chemie,* **70** (1966) 883
233. D. Clotman, D. Van Lerberghe, Th. Zeegers-Hyuskens, *Spectrochim. Acta A,* **26** (1970) 1621
234. J. P. Muller, G. Vercruysse, Th. Zeegers-Hyuskens, *J. Chem. Phys. Phys. Chim.,* (1972) 1439
235. W. Caminati, *J. Mol. Spectrosc.*, **86** (1981) 193
236. S. Vazquez, R. A. Mosquera, M. A. Rios, C. van Alsenoy, *J. Mol. Struct. (Theochem),* **184** (1989) 323
237. S. Vazquez, R. A. Mosquera, M. A. Rios, C. van Alsenoy, *J. Mol. Struct., (Theochem),* **181** (1988) 149
238. R. Friedemann, A. Jabs, *J. Mol. Struct. (Theochem.),* **283** (1993) 191
239. S. Hobohm, Master thesis Martin-Luther-University (1996)
240. R. Friedemann, A. Fengler, S. Naumann, U. Gromann, *J. Mol. Struct. (Theochem.),* **357** (1995) 217
241. J. F. Bacon, J. H. van der Maas, J. R. Dixon, W. O. George, P. S. McIntyre, *Spectrochim. Acta,* **45** (1989) 1313
242. L. P. Kuhn, *J. Am. Chem. Soc.*, **76** (1954) 4323
243. S. Hwang, N. Tamilarasu, K. Ryan, I. Huq, S. Richter, W. Clark Still, T. M. Rana, *Proc. Nat. Am. Soc*., **96** (1999) 12997
244. R. M. Badger, S. H. Bauer, *J. Chem. Phys.*, **5** (1939) 839-841
245. F. Schuppe, I. Schaller, A. Kolbe, I. Wawer, *J. Mol. Liqu.*, **34** (1987) 223
246. A. Kolbe, P. Kaps, M. Plass, *Zeitschrift für Physikal. Chemie,* **191** (1995) 191
247. A. Gröbel, Ph. D. Thesis, Martin-Luther-University Halle-Wittenberg (1998)
248. M. Carisma, F. Formaggio, G. Valle, C. Toniolo, M. Saviano, R. Iacovino, L. Zaccaro, E. Benedetti, *Biopolymers*, **42** (1997) 1
249. O. Pieroni, A. Fissi, C. Pratesi, P. A. Temussi, F. Ciardelli, *Biopolymers*, **33 (**1993) 1
250. N. Hirota, K. Mizuno, Y. Goto, *Protein Sci.*, 6 (1997) 416
251. D. S. Wishat, B. D. Sykes, F. M. Richards, *J. Mol. Biol.*, **222** (19991) 311
252. A. Perczel, I. Lengyel, H. H. Mantsch, G. D. Fasman, *J. Mol. Struct.*, **297** (1993) 115

253. C. Benamon, L. Bellon, *C. R. Acad. Sc. Paris*, **393** (1981) 425
254. A. Stimac, J. Kobe, *Synthesis*, (1990) 461
255. A. Temeriusz, B. Piekarska-Bartoszewicz, I. Wawer, *Carbohydr. Res.*, **304** (1997) .335
256. B. Piekarska-Bartoszewicz, A. Temeriusz, *J. Carbohydr. Chem.* **12** (1993). 913
257. M. H. Abraham, *Pure Appl. Chem,*. **65** (1993) 2503
258. Public domain, http:\www.acs.
259. J. K. Kauppinen, D. J. Moffatt, H. H. Mantsch, D. G. Cameron, *Appl. Spectr.*, **35** (1981) 271
260. J. K. Kauppinen, D. J. Moffatt, H. H. Mantsch, D. G. Cameron, *Analyt. Chem.*, **53** (1981) 1454
261. J. K. Kauppinen, D. J. Moffatt, D. G. Cameron, H. H. Mantsch, *Appl. Optics,* **10** (1981) 1866
262. M. A. Czarnecki, Y. Ozaki, *Spectrochim Acta*, **52** (1996) 1593
263. F. Long, T. B. Freedman, T. J. Tague, L. A. Nafie, *Appl. Spectrosc*., **51** (1997) 508
264. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. P. Stewart, *J. Am. Chem. Soc.*, **107** (1985) 3902
265. F. J. Devlin, P. J. Stephens, J. R. Cheeseman, M. J. Frisch, *J. Am. Chem. Soc*., **118** (1996) 6327
266. P. J. Stephens, *J. Phys. Chem*., **89** (1985) 748
267. P. J. Stephens, *J. Phys. Chem*., **91** (1987) 1712
268. J. S. Anderson, K. Saddington, *J. Chem. Soc.*, (1949) 381
269. C. Benamon, L. Bellon, *C. R. Acad. Sc. Paris,* **393** (1981) 425
270. M. Plass, Master thesis, Martin-Luther-University Halle-Wittenberg (1990)

E APPENDIX

E.1 Experimental

E.1.1 Materials

E.1.1.1 Guanine derivatives

The guanine derivatives ACV (9-(2-hydroxyethoxymethyl)guanine), NAcACV (N^2-acetylacyclovir), OAcACV (*O*-acetylacyclovir), diAcACV (N^2, *O*-diacetylacyclovir), DCV (2-amino-9-(2-hydroxyethoxymethyl)-9H-purine) and OAcDCV (*O*-acetyldeoxyacyclovir) were synthesised at the National Institute of Chemistry (Ljubljana, Slovenia) as described[254] and kindly donated for this study. They were stored in dry atmosphere in order to neglect water traces which could influence the tautomeric equilibrium.

E.1.1.2 Amino acid and oligopeptide derivatives

The amino acid and peptide derivatives, Ac-Ala-OMe, Ac-Phe-OMe, Ac-Val-OMe, Ac-Ala-Ala-OMe, Ac-Ala-Ala-Ala-OMe, Ac-Ala-Ala-Ala-Ala-OMe, were purchased from BACHEM Biochemica (Heidelberg). Their enantiomeric purity was higher than 99 %. N-methyl acetamide was received from ALDRICH. The substances were stored in a refrigerator and dried before use in a desiccator filled with molecular sieve.

The tripeptide and dipeptide derivatives were synthesised according to the HOBt/DCC method in the group of C. Griehl (Department of Biotechnology and Applied Food, Technical University Köthen); The purity of the compounds was tested chromatographically and exceeded a diastereoisomeric purity of 95%. Other physical characteristics are described elsewhere[165]

The functional groups at the N- and C- terminal end were blocked by protecting groups of the peptide chemistry. The Z (benzyloxycarbonyl), Ac (acetyl) and the Boc (tert. butyloxycarbonyl) residues were used for the protection of the N- terminal site. The acidic function was blocked by ester residues which were either methyl (OMe), ethyl (OEt), tert. butyl (OtBu) or benzyl (OBzl) esters. Polar side chain groups such as OH and NH_2 functions were in selected cases protected by methoxy, benzoxy or benzylamine residues.

E.1.1.3 Ureido sugar derivatives

The compounds used in this study were synthesised in the group of A. Temeriusz, Department of Chemistry, Warsaw University, according to a procedure described elsewhere[255, 256]. They were dried over molecular sieves and dissolved in methanol at a concentration of 0.01 mol/l for the HPLC measurements. The chromatogram of the substances did not show an indication for impurities higher than 1%.

E.1.1.4 Phospholipids

As models for the phospolipid membrane 1.2-dihexadecyl-sn-glycero-3-phosphoryl choline (DHPC) and 1.2-dimyristoyl-sn-glycero-3-phosphoryl choline (DMPC) were purchased from BACHEM Biochemica. 1.2-dipalmitoyl sn-glycero-3-phospho choline (DPPC) was available from FLUKA as its monohydrate.

E.1.1.5 Alcohols and diols

The monomeric alcohols propan-1-ol (RIEDEL DE HAEN), propan-2-ol (MERCK) and n-hexan-1-ol (REACHIM) were purchased in the highest possible grade (spectr. grade) and stored over molecular sieves. Furthermore, the diols propan-1.2-diol (FLUKA), propan-1.3-diol, butan-1.4-diol, pentan-1.5-diol (MERCK, pro analysis) and hexan-1.6-diol were used in this study. As cyclic proton donor compounds with two functional OH groups cis-cyclohexan-1.2-diol and trans- cyclohexan-1.2-diol and the racemic cyclohexan-1.4-diol purchased from FLUKA were tested.

The proton acceptor compounds were pyridine, 2.4-lutidine and 2.4.6-collidine were freshly distilled under vacuum immediately before use.

The hydrogen bonding behaviour of amino acid, dipeptide and tripeptide derivatives with alcohols was investigated in methylene chloride solutions. The alcohols were 1.1.1.3.3.3-hexafluoropropanol (MERCK, spectr. grade) and phenol (FLUKA).

E.1.1.6 Solvents

For the HPLC measurements the standard organic phase was acetonitrile (RATHBURN, Walkerburn, UK)). The aqueous phase was a 0.05 molar ammonium acetate buffer solution of a pH of 7.4 (FISON, Loughborough, UK). In selected experiments the HPLC solvents methanol and 1.1.1-trifluoroethanol (FISHER CHEMICALS, Loughborough, UK) were used as organic modifier.

Solvents in the NMR spectroscopic experiments were D_2O (ALDRICH), $CDCl_3$ (MERCK) and DMSO-d6 (ALDRICH) of the highest available grade. The percentage of water in sealed bottle of DMSO-d6 was less than 0.01%.

Solvents for IR and UV/VIS spectroscopy were methylene chloride (FLUKA), acetonitrile (FLUKA, 190 super purity solvent, ROMIL pure chemistry), carbon tetrachloride (ALDRICH) were kept over molecular sieves.

E.1.2 HPLC measurement

A Hewlett-Packard 1090 series high performance liquid chromatograph was used. Data acquisition and processing were performed on a Viglen IBM compatible PC with HP Chemstation software (HEWLETT-PACKARD, Amsterdam, The Netherlands). The system was set up with a low pressure pump and an autosampler. The solutions (3, 5 or 10µl) were injected onto the HPLC system twice. The injected volume varied according to the sample's property to absorb in the UV region of the array detector.

According to their solubility ca. 1mg of the sample was dissolved in a 1:1 mixture of AcN/buffer. More polar compounds such as the ureido sugar derivatives and peptides with polar side chain residues were dissolved in methanol. The concentration of the peptide solution was ca. 0.005 to 0.01 mol/l. As a standard set-up the signals at 215, 230 and 245 nm were monitored. The average retention time was used for further analysis.

E.1.2.1 Test compounds for the characterisation of the chromatographic systems

The 30 test compounds given in table E1 are used for the calibration of the chromatographic system. For these compounds the solute descriptors are known. They were selected to represent a wide variety of solute properties from the UCL database. The bold written compounds built the standard set which were injected as a mixture for the calibration of the CHI value.

Table E1: Test compounds for the determination of the coefficients and their descriptors[74]

Name	R_2	π_2^H	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x
Paracetamol	**1.06**	**1.78**	**1.09**	**0.81**	**1.172**
Acetanilide	**0.87**	**1.40**	**0.50**	**0.67**	**1.137**
Acetophenone	**0.82**	**1.01**	**0.00**	**0.48**	**1.014**
Propiophenone	**0.80**	**0.95**	**0.00**	**0.51**	**1.155**
Butyrophenone	**0.80**	**0.95**	**0.00**	**0.51**	**1.296**
Valerophenone	**0.80**	**0.95**	**0.00**	**0.50**	**1.437**
Hexanophenone	**0.72**	**0.95**	**0.00**	**0.50**	**1.578**
Heptanophenone	**0.72**	**0.95**	**0.00**	**0.50**	**1.718**
Octanophenone	**0.72**	**0.95**	**0.00**	**0.50**	**1.859**
4-Iodophenone	1.38	1.22	0.68	0.20	1.033
Dibenzothiophene	1.96	1.31	0.00	0.20	1.379
4-Chlorophenol	0.92	1.08	0.67	0.20	0.898
4-Cyanophenol	0.94	1.63	0.80	0.29	0.930
Benzamide	0.99	1.5	0.49	0.67	0.973
Caffeine	1.50	1.6	0.00	1.33	1.363
Indasole	1.20	1.22	0.53	0.35	0.905
Anisole	0.71	0.75	0.00	0.29	0.916
Benzonitrile	0.74	1.11	0.00	0.33	0.871
Chlorobenzene	0.72	0.65	0.00	0.07	0.839
Naphthalene	1.34	0.92	0.00	0.20	1.085
Dinitrobenzene	1.13	1.63	0.00	0.46	1.065
Phenol	0.81	0.89	0.60	0.30	0.775
Trifluoromethylphenol	0.43	0.87	0.72	0.09	0.969
Toluene	0.60	0.52	0.00	0.14	0.857
Corticosterone	1.86	3.43	0.40	1.63	2.739
Aniline	0.96	0.96	0.26	0.41	0.816
Testosterone	1.54	2.59	0.32	1.19	2.383
Hydrocortisone-21-acetate	1.89	3.67	0.43	1.90	3.095
p-Toluidine	0.92	0.95	0.23	0.45	0.957
m-Nitroaniline	1.20	1.71	0.40	0.35	0.990

E.1.2.2 Chromatographic columns

Table E2 and E3 give the parameters of all chromatographic systems tested in the course of our investigations. For the determination of the CHI values finally the five systems in table E2 were chosen.

Table E2: HPLC columns used for the isocratic measurements and for the descriptor determination

Name	Abbreviation	Particle size	Column dimension	Supplier
ODS2-IK5 Inertsil	In	5 µm	150x4.6 mm	A
Prodigy ODS2	Pro	5 µm	150x4.6 mm	D
Novapak-CN	CN	4 µm	75x3.9 mm	B
RexChrom IAM PC2	IAM	12 µm	150x4.6 mm	H
Nucleodex β PM	CD	5 µm	200x2.6 mm	H

Supplier:
A: CAPITAL HPLC, Broxburn, UK.
B: WATERS CHROMATOGRAPHY, Watford, UK
C: SUPELCO, UK.
D: PHENOMENEX UK, Macclesfield, UK.
E: PHASE SEPARATIONS, Deeside, UK
F: BIOTAGE UK, Division of Dyax Hertford, UK
G: POLEMER LABORATORIES, Separation Science Division, Stratton, UK.
H: FISHER SCHIENTIFIC UK, Loughborough, UK

Table E3: HPLC columns for the selection of appropriate coefficients for further studies

Name	Abbreviation	Particle size	Column dimension	Supplier[a]
Luna ODS	Luna1	5 µm	150x4.60mm	D
Luna ODS	Luna2	5 µm	50x4.60mm	D
Symmetry C18	SY	3.5 µm	50x4.6 mm	B
NovaPak RP	NRP	4 µm	75x3.9 mm	B
Supelcosil ABZ	ABZ	3 µm	33x4.6 mm	C
Selectosil RP	SRP	5 µm	150x4.6 mm	D
Sherisorb ODS1	OD1	5 µm	150x4.6 mm	E
Unisphere-C18	BRP	8 µm	100x4.6 mm	F
Unisphere PBD	BPB	8 µm	100x4.6 mm	F
Asahipak ODP	APO	5 µm	150x4.6 mm	D
RPS 40	POL	8 µm	50x4.6 mm	G
Novapak Phenyl	NPH	4 µm	75x3.9 mm	B
Diol-YC5 Inertsil	Indiol	5 µm	150x4.6 mm	A
Selectosil-diol	Sdiol	5 µm	150x4.6 mm	D
Nucleosil NH2	NH2	5 µm	150x4.6 mm	F

[a]For supplier code see above

E.1.2.3 Isocratic Measurements

The determination of φ_0 is based on isocratic measurements which was done for the compounds of the standard set (bold written compounds in table E1) and selected amino acid and peptide derivatives on the chromatographic systems listed in table E2. For these columns a low rate of 1.5 ml/min. can be realised except for the permethylated β- cyclodextrine phase (CD) which worked at a flow rate was 1ml/min.

In the isocratic experiment the concentration on acetonitrile was increased in 5 % steps.

The experiment started with a calibration time of 5 min. to allow an adjustment of the column to the solvent mixture. The, the sample was injected and the chromatogram was recorded over 25 min. at constant solvent mixture. A washing time of 5 min with pure acetonitrile guaranteed that the sample was completely removed from the stationary phase before a new cycle was started.

For these measurements the temperature was kept constant at 25 °C.

The retention factor k can be calculated according to eq. E1:

$$k = \frac{t_{R1} - t_0}{t_0} \qquad [E1]$$

Plotting logk vs. the concentration of the organic modifier gives a function which allows a linear regression analysis. The intercept of the regression line is defined as $logk_0$. $Logk_0$ and the slope S are needed to calculate φ_0 according to equation 6.

Temperature dependent measurements were performed on the Luna ODS column (Luna1) at 25 and 40 °C which were kept constant with the help of a column thermostat at a flow rate of 1 ml/min.

E.1.2.4 Linear gradient measurements

E.1.2.4.1 Variation of the gradient time

To test the influence of the gradient time on the curving different gradient measurements were performed using the standard set and the ODS2-IK (In) column. The following experiments were performed:

A)	0.0-0.5 min.	0% AcN
	0.5-2.5 min.	100% AcN
	2.5-3.5 min.	100% AcN
	3.5-3.7 min.	0 % AcN
	3.7-5.0 min	0 % AcN
B)	0.0-0.5 min.	0% AcN
	0.5-4.0 min.	100% AcN
	4.0-5.0 min.	100% AcN
	5.0-5.2 min.	0 % AcN
	5.2-7.5 min	0 % AcN

C)	0.0-5.0 min.	0% AcN
	5.0-30.5 min.	100% AcN
	30.5-35.0 min.	100% AcN
	35.0-40 min.	0 % AcN
	40-45 min	0 % AcN

The correlation factors show that the best correlation was received from experiment B which was chosen for the CHI determination.

Table E4: Test of the influence of the gradient time t_G on the curving (Inertsil)

Gradient experiment	r^2	Slope A	Intercept B
A)	0.9612	49.645	-107.47
B)	0.9843	31.105	-73.231
C)	0.9715	4.8457	-15.497

E.1.2.4.2 Effect of the linear gradient on the bandwidth

A linear gradient elution was performed varying the composition of the organic modifier from 0 to 100 % over a gradient time of 2.5, 3.5, 5.5, 7.5, 10, 15, 30, 45, 60 and 90 min. The column temperature was kept constant at 25 °C (±0.5 K). Thus, the retention times and bandwidths vary less than 1% and 5%, respectively. The bandwidth of the signals was determined automatically with the help of the HP ChemStation software.

E.1.2.4.3 Effect of organic modifier

Different solvents were used as organic modifiers; such as AcN, MeOH and TFE. In all experiments the aqueous phase was 0.05 molar ammonium acetate buffer (pH 7.4).

A short ODS column (Luna2) allowed a flow rate of 2 ml/min still giving well resolved signals for the standard set. Using a gradient time of 2.5 min. leads to similar results as received for the Luna1 column. The retention times used for the plot in Fig. 11 are given in table A6.

E.1.2.5 Calculation of CHI

The calculation of the CHI values for all compounds is based on the correlation of the retention times t_R measured under linear gradient conditions and the isocratically determined φ_0 of the standard set (table E2)[70]. Performing a linear regression of φ_0 vs. t_R of the compounds of the standard set gives the parameters A and B in eq. E2:

$$\varphi_0 = A \cdot t_{RS} + B \qquad [E2]$$

These parameters were then used to calculate the CHI values for the test compounds or for the peptide derivatives according to equation E3.

$$CHI = A \cdot t_{Rx} + B \qquad [E3]$$

Hence, the CHI value should be similar in the magnitude as φ_0 but it is measured under gradient conditions.

E.1.2.6 Determination of the coefficients characterising the chromatographic partitioning system

Based on this correlation the CHI values for the 30 test compounds (table E1) with known descriptors were determined and are given in table A5. The multiple regression analysis carried out using the Drugidea TM software package (CHEMICRO Ltd., Budapest, Hungary) and the SmartWare II (INFORMIX Software Inc.) gives the coefficients c, r, s, a, b and ν (table 5) which characterise the partitioning system. The determination of the CHI values for the 30 test compounds was done for all chromatographic systems listed in table E2 and E3. Based on the variety of the column parameters the five columns in table E2 were selected for the peptide studies.

The CHI values for peptide derivatives are in part discussed in the results. Further data necessary for the deduction of fragmental descriptors are given in table A7.

The CHI values of the ureido sugars are given table A10.

E.1.2.7 Determination of solute descriptors

V_x and R_2 were calculated with a self-made software routine[257]. The calculation of the excess molar refraction R_2 is based on the refractive index which is an additive property and can be approximated from substructure using the ACD/Chem Sketch software[258].

With known values for V_x and R_2 the equation system eq. 7 is reduced to three unknown parameters. Using the solver function in EXCEL (MICROSOFT OFFICE 97) the best fitting results for π_{2H}, $\sum\alpha_2^H$ and $\sum\beta_2^H$ can be found.

The descriptors for the peptide derivatives are discussed in the main part of the thesis. Additional experimental data are given in table A8 and A9.

The experimentally determined descriptors of the ureido sugar derivatives can be found in table A10.

E.1.2.8 Calculation of the fragmental CHI values and fragmental descriptors

Based on the additivity concept the CHI value or the descriptor of a tripeptide of the general formula X-^{1}Xaa-^{2}Xaa-^{3}Xaa-Y should be the sum of the contributions of the fragmental data of X, ^{1}Xaa, ^{2}Xaa, ^{3}Xaa and Y. Hence, we can built an equation system for all investigated tripeptides which will find the minimum of the least square sum with the help of the solver function in EXCEL (MICROSOFT OFFICE 97).

E.1.3 Infrared measurements

With the exception of the phospholipid-peptide mixtures the studies were performed on the FTIR spectrometers IFS 25 and IFS 66 (BRUKER). The irradiation was produced by globar; a traditional DTGS was used as a detector. The interferogram was transformed by phase correction, zero filling factor 2. The sample spectra were recorded in the absorbance mode using the pure solvent as a background. In general the resolution was 2 cm^{-1}. In selected cases the NH region of the ureido sugars the resolution was enhanced on 0.25 cm^{-1} which can be realised on the IFS 66. The number of the accumulated scans varied between 32 and 100 but was constant within a group of measurements. To reduce the H_2O concentration in the instruments they were permanently purged with dry air.

The thickness of the IR cell and concentration was adapted to the absorption characteristic of each sample solutions.

E.1.3.1 Temperature measurements

Usually the measurements were performed at 25 °C. For the intermolecular association of alcohols and diols the temperature was varied in the range of 5 to 30 °C which was realised by a sample holder purged with thermostated water received from a cryostat. The accuracy of the temperature can be given in ±0.2 K in the sample cell which was permanently controlled by an external thermocouple.

The intramolecular hydrogen bond behaviour was tested on selected substances between 0 and 50 °C and -15 to 30 °C for the ureido sugar derivatives and the diols, respectively, using the variable temperature cell P/N 21500 (GRASEBY SPECAC). The cooling agent was a methanol/dry ice mixture. The accuracy of the temperature was checked by a separate thermocouple and did not exceed the chosen temperature by more than ± 0.2 K. The temperature was varied in 5 K steps with a 5 min delay before the measurement in order to allow the temperature adjustment. The measurements were performed in a 1 mm NaCl cell or an CaF_2 cell for low temperature studies.

The temperature variations for the study of the intramolecular association behaviour of the peptide derivatives were realised using a temperature device equipped with Peltier cooling elements. Spectra at -10, -5, 15, 25 and 40 °C were recorded in solutions of carbon tetrachloride using a 3mm NaCl cell.

E.1.3.2 Solvent effect on the conformers

E.1.3.2.1 Peptide investigation

Methylene chloride solutions of peptide concentration of 0.003 to 0.009 mol/l are sufficiently transparent in the NH and amide I region also when a NaCl cell of 3mm thickness was used. The solutions of the polar solvents (HFiP, D_2O and DMSO-d6) were performed in a CaF_2 cell (0.114 mm). The concentration was accordingly higher and varied between 0.03 and 0.09 mol/l.

Some dipeptide and tripeptide derivatives were found to be sufficiently soluble in apolar solvents as carbon tetrachloride. At concentrations of 0.002 for the dipeptide and 0.005 mol/l for the tripeptide derivatives no intermolecular hydrogen bonds were observed.

The region of the ester carbonyl and the amide I and II vibration was cut from the whole spectra. A two-point base line correction was applied.

E.1.3.2.1.1 Phospholipid:peptide mixtures

Measurements were performed on a Vector 22 FTIR spectrometer (BRUKER) using a thermostated out-of-compartment overhead ATR unit. The heating and cooling curves were controlled by the instrument's computer. Ca. 0.7 mg of the compounds were dissolved in 100µl of a mixture of $CHCl_3$:MeOH (1:1). In the mixtures of DHPC and the tripeptide derivative the stochiometric amount of the peptide was added to the phospholipid solution. Then, the reaction vessel was sonicated for 10 minutes at a temperature above the phase transition temperature of the lipid. The solution was slowly spread on a trapezoid-shaped ZnSe-ATR crystal (dimensions: 72x10x6mm, face angle ϑ=45°) which was mounted in the ATR unit (GRASEBY SPECAC). The lower part of the compartment could be cooled or heated by an agent continuously flowing from a cryostat which allowed a temperature variation between -14 and 85 °C. A separate thermocouple showed that the temperature in the crystal was constant after a delay of 3 min. tolerating a temperature gradient of ±0.1 K. The substance multilayer on the crystal surface was always higher than the penetration depth of the evanescent wave. Then, the upper side of the ATR plate was covered with a self-built cell

which allowed purging the crystal with dry nitrogen or covering the material with ca. 0.5 ml D_2O in order to guarantee full hydration.

E.1.3.2.2 Ureido sugar derivatives

Infrared spectra of different ureido sugars were recorded as $CHCl_3$ solutions in concentrations of ca. 0.001 mol/l. A NaCl cell with a thickness of 3mm was used for all investigations.

E.1.3.3 Intermolecular association behaviour

E.1.3.3.1 Alcohol-amine associates

A CaF_2 cell of the thickness 3mm was used for these studies. The alcohol solutions in methylene chloride were sufficiently diluted (ca 0.03 mol/l) in order to avoid intermolecular association. In order to form the associate complex the proton acceptor compounds were added to the alcohol solution. Hence, the amine concentration varied between 0.02 and 1.0 mol/l. Particular attention was taken in the region of low amine concentration by checking the concentration by an internal signal of the pyridines.

The alcohol and the amine complexes were investigated at 12 and 32 °C which was realised in a water cooled sample holder purged with thermostated water.

The procedure was as follows: At first the alcohol solution was recorded at both temperatures. Then, the amine was added and the measurements were repeated beginning at low temperatures.

E.1.3.3.2 Alcohol-peptide associates

Investigations at 25 °C were performed using a 3mm NaCl. The procedure of the measurements was similar as described for alcohol-amine associates: The alcohol dissolved in methylene chloride (0.01- 0.005 mol/l) were recorded without proton acceptor compounds and with different amounts of the peptide compound. The peptide concentration varied in the range of 0.005 to 0.05 mol/l. As a measure for the association the integral absorbance between 3655 and 3480 cm^{-1}, and 3635 and 3525 cm^{-1}, for HFiP and phenol, respectively, was taken as a measure for the free OH species. Furthermore, the signals of the acceptor functions were used to calculate the equilibrium constant after a peakfitting procedure was applied to the complex band profile.

E.1.3.4 Band deconvolution and peakfitting

There are two general approaches in the analysis of complex band profiles namely the band deconvolution and the peakfitting method.

The identification of the number of overlapping bands can be done with the deconvolution method implemented in the OPUS software package (BRUKER) which narrows the band width using Lorentzian band profiles. In the literature the difficulties of this technique are widely discussed[259, 260, 261, 262]. On the example of Z-Val-Val-Val-OMe we tested the influence of the deconvolution factor on the ratio of the integral intensity of the bands always keeping in mind that no overcompensation occurs.

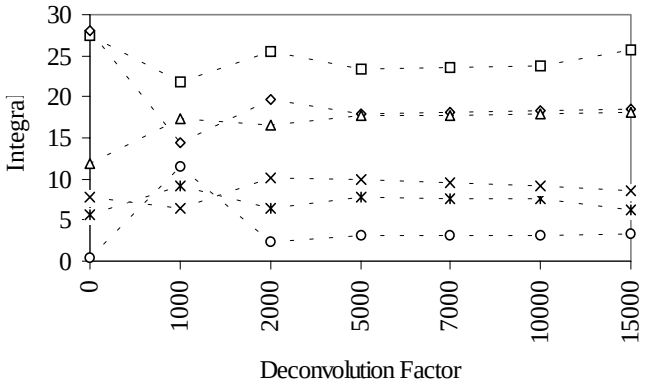


Fig. E3: Effect of the deconvolution factor on the ratio of the integral intensities of the deconvoluted bands of Z-Val-Val-Val-OMe in the region between 1800 and 1600 cm^{-1}

At a deconvolution factor of 0 we have plotted the integrals received from peakfitting of the original spectra. Our tests showed that the intensity ratios of the fitted signals is the same as the ratio of the deconvoluted signals of the pure tripeptide solutions.

Hence for a quantitative analysis we have taken the number of signals from the deconvolution using a deconvolution factor of 2000 and applied a peakfitting procedure on the original data in the baseline corrected IR region between 1800 and 1600 cm^{-1}.

Here we have to chose between two different software packages: the peakfit function implemented in the OPUS software allowing a simulation with Gauss-Lorentzian sum functions and the „Peakfit-program for non linear curve fitting“, Vs. 4.0 and Vs. 5.0 (JANDEL SCIENTIFIC) which has a higher variety of functions to simulate IR band profiles. With the exception of the studies of the intramolecular association behaviour of the peptide derivatives (Pearson function) and the ureido sugar derivatives (Voigt functions), Gauss-Lorentzian sum function were used in the fitting procedure. A Levenberg-Marquardt procedure was followed to receive a coincidence between experimental and fitted band profile with $r^2 \geq 0.9995$. The band maxima and the integral areas were taken as a measure in the discussion.

In order to study the intermolecular association the acceptor compounds were recorded at different concentrations. The peakfitting approach was applied only allowing the band intensity to vary. The parameters of the fitted Gauss-Lorentzian sum functions are given in table E5.

The integral intensities were plotted vs. the concentration. The linear regression of the curves of gives the integral molar extinction coefficients. In the presence of the alcohol two additional signals were introduced but for the non-associated signals the band profile was kept unchanged. (Table E5)

Table E5: Parameters for the fitting procedure (the parameters of the associate bands are not given)

Compound	Peak position	Bandwidth	Band shape (in % Lorentzian)
Ac-Ala-OMe	1742.4	21.14	24
	1678.9	19.38	68
Ac-Phe-OMe	1743.1	20.29	30
	1679.6	20.92	88
Ac-Val-OMe	1739.0	19.43	18
	1680.6	18.76	66
Z-Ala-Ala-OMe	1742.3	21.6	43
	1720.4	26.68	35
	1683.9	24.34	68
Z-Phe-Phe-OMe	1743.9	20.4	43
	1722.8	28.5	21
	1682.7	26.36	52
Z-Val-Val-OMe	1739.7	20.46	39
	1721.2	24.3	5
	1682.7	22.8	61
Z-Ala-Val-OMe	1740.2	18.47	22
	1721.8	35.88	42
	1685.8	23.13	71
Z-Ala-Ala-Ala-OMe	1742.8	22.81	34
	1720.0	27.00	27
	1689.4	23.39	39
	1671.2	23.73	27
Z-Phe-Phe-Phe-OMe	1744.4	21.51	21
	1722.1	28.61	28
	1685.7	29.56	29
	1669.2	22.73	22
Z-Val-Val-Val-OMe	1740.3	19.29	15
	1721.1	24.66	42
	1690.5	22.88	33
	1672.8	25.24	44
Z-Ala-Val-Val-OMe	1740.6	20.62	27
	1720.8	31.94	0
	1689.4	20.42	29
	1671.8	25.79	36

Table E6: Integral molar extinction coefficients received from linear regression analysis of the fitted spectra (1800-1600 cm^{-1}) and from the integral of the NH band [in l/mol*mm]

Substance	ν_{NH}		ν_{Ester}		$\nu_{Ac/Z}$		$\nu_{PeptideI}$		$\nu_{PeptideII}$	
	$\varepsilon_{int.}$	r^2	$\varepsilon_{int.}$	r^2	$\varepsilon_{int.}$	r^2	$\varepsilon_{int.}$	r^2	$\varepsilon_{int.}$	r^2
Ac-Ala-OMe	358	0.9923	1028	0.9807	1631	0.9799				
Ac-Val-OMe	356	0.9927	1018	0.9945	1722	0.9945				
Ac-Phe-OMe	372	0.9923	1055	0.9914	1885	0.9921				
Z-Ala-Ala-OMe	664	0.9904	1425	0.9748	1503	0.9717	1895	0.9883		
Z-Val-Val-OMe	730	0.9966	1333	0.9963	1206	0.9971	1800	0.9978		
Z-Ala-Val-OMe	680	0.997	801	0.995	2043	0.9978	1636	0.9949		
Z-Phe-Phe-OMe	775	0.9819	1181	0.9916	1467	0.9936	1782	0.9819		
Z-Ala-Ala-Ala-OMe	826	0.9971	1212	0.9971	1521	0.9963	1296	0.9970	1715	0.9930
Z-Val-Val-Val-OMe	829	0.9942	909	0.9827	1478	0.984	533	0.9474	2193	0.987
Z-Ala-Val-Val-OMe	887	0.9936	959	0.9914	1553	0.9918	1056	0.9845	1856	0.9984
Z-Phe-Phe-Phe-OMe	960	0.9556	1093	0.9666	1408	0.9693	1623	0.9687	1182	0.9556

E.1.4 NIR measurements

Our IFS 66 can be adapted to the near infrared region. For the measurement of the overtone signal of the NH function the globar is still delivers sufficient irradiation which can be detected with the DTGS detector. For the absorbance spectra 100 scans were accumulated. The solutions were measured in a 2 cm quartz cell using the pure solvent as background spectra.

E.1.5 VCD measurements

The measurements and calculations were performed in the group of L. A. Nafie and T. B. Freedman at the University of Syracuse. Solutions of the diastereoisomers of Z-Ala-Phe-Val-OMe in carbon tetrachloride (ALDRICH, reagent grade, stored over molecular sieves) were prepared at concentrations of ca. 0.002 mol/l. The solutions were recorded in the CO stretching region using a BaF_2 cell with 5 mm spacers. All studies were performed at -5, 5 and 25 °C. The IR and VCD spectra were recorded on a NICOLET MAGNA 850 FTIR spectrometer equipped with an accessory VCD bench at a resolution of 4 cm^{-1}[263]. IR measurements consisted of 32 scans. VCD measurements were collected in simultaneous rapid-scan mode (10 000 scans, 3.75 hours) and have been corrected for baseline artefacts by subtraction of solvent VCD spectra.

E.1.6 Quantum mechanical calculations

Two general quantum mechanical approaches were applied for the geometry optimisation of molecules:

- Semiempirical approach (AM1) uses on standard parameterisation of bond length and angles. The Austin Model (AM1)[264] has given reliable results for the optimisation of compounds with intramolecular hydrogen bonds. The AM1 method is implemented in the HYPRCHEM package (Vs. 4.5) which was used for the geometry optimisation of the guanine derivatives and the ureido sugar derivatives.
- Ab initio calculations were carried out on the model compound MeOCO-Ala-Ala-NMe (LL) using the Hartree-Fock 6-31G(d) level of Gaussian 98 (Gaussian Inc. Pittsburgh, PA). These calculations were performed by T. B. Freedman at the University of Syracuse and were the requirement for the calculation of VCD spectra using magnetic field perturbation calculations (MFP) of the VCD intensities[265, 266, 267].

E.1.6.1 Guanine derivatives

Due to the ethoxy chain the number of conformers received by rotation of this function would be relatively large. Thus, to start with we have taken the structures which were known to be preferred from previous results[44] of a force field calculation. In fact, backfolded conformation was the most stable arrangement. A number of starting conformations was created recognising the possible occurrence of intramolecular hydrogen bond formation based on the backfolded conformer. As a result the differences of the total energies and the net charges of the atoms were discussed in comparison to the NMR results.

E.1.6.2 Ureido sugar derivatives

The energy optimisation of the N-(2-amino-2-deoxy-β-D-glucopyranoside)-N'-carbamoyl-L-amino acid and dipeptide esters was performed with the help of the semi-empirical AM1 calculation method implemented in the HYPERCHEM program package. Therefore, a number of initial conformations was created recognising the occurrence of possible intramolecular hydrogen bonds. The energy optimisation was performed without any restrictions regarding dihedral angles.

E.1.7 UV/VIS measurements

A double beam UV/VIS spectrometer Lambda 16 (PERKIN ELMER) was used for the recording of the UV/VIS spectra in the region between 500 and 190 nm. A slit width of 2 nm and recording speed of 120 nm/min was used as experimental parameters. Solutions of the compounds in H_2O, methylene chloride and acetonitrile solutions were measured in a 1cm quartz cell using the pure solvents as reference probes. The concentrations were approximately 10^{-6} mol/l.

E.1.8 CD measurements

Using a J720 spectropolarimeter (JASCO) the solutions of Ac-Val-OMe, Z-Val-Val-OMe, Z-Val-Val-Val-OMe and Z-Val-D-Val-Val-OMe were investigated between 180 and 260 nm in different solutions and in dependence of the temperature. The cell length of the quartz cell (QS, HELLMA) was 0.0096 cm. At a resolution of 0.1 nm and a band width of 1.0 nm the spectra were scanned with a speed of 10 nm/min. The sensitivity of the detector was 20 mdeg. Four scans were accumulated for each spectrum. Measurements of the AcN and HFiP solutions were performed at 10, 20, 30 and 40 °C. Due to the absorbance of the solvent in the far UV the methylene chloride solutions were recorded in the region between 210 and 260 nm at 10, 15, 20 and 25 °C. The temperature in the cell which was put into a thermostated sample holder was checked with a external thermocouple. The temperature accuracy was ±0.2 K. The concentration of the solution was ca. 0.005 mol/l in order to guarantee similar condition as in the IR spectra. It was interesting to study the effect of small amounts of HFiP on the conformational behaviour of the peptides. Therefore, the compounds were dissolved in acetonitrile and methylene chloride solutions containing ca. 0.01 to 0.005 mol/l HFiP. In these cases the HFiP solutions were recorded as background spectra, otherwise the pure solvents were used.

E.1.9 NMR-measurements

E.1.9.1 Standard measurements

^{1}H NMR spectra of the saturated D_2O solutions were recorded on a 500 MHz and on a 400 MHz NMR spectrometer (VARIAN). 32 scans were accumulated for the ^{1}H NMR spectra at a temperature of 27 °C for routine measurements, Otherwise, the number of scans was chosen according to the sensitivity of the sample. In selected cases ^{1}H/^{13}C correlation experiments (COSY and NOESY) of the DMSO-d6 solutions were performed to support the assignment of the signals.

E.1.9.2 Temperature measurements

^{1}H NMR spectra of the ureido sugar derivatives were recorded on a VARIAN Unity Plus 500 MHz spectrometer equipped with a variable temperature probe. The concentration of the ureido glucoside derivatives dissolved in $CDCl_3$ and in DMSO-d6 was chosen similarly to the IR experiment. The temperature was varied in the range between 25 and 60 °C in 5 K steps.

E.1.9.3 H/D exchange experiements

The H/D exchange measurements were performed in $CDCl_3$ and followed the procedure described. During the experiment the temperature was kept constant at 25 °C.

E.1.9.4 HFiP titration

The peptide derivatives were dissolved in $CDCl_3$ at concentrations of 0.008 to 0.01 mol/l. 1ml of the solution was filled in a 5 mm NMR tube. The recording of the ^{1}H NMR signals was done on a 400 MHz NMR spectrometer (VARIAN). Generally, 32 scans were

accumulated at a temperature of 27 °C. Then, defined amount of HFiP were added to the solution in the NMR tube. The carefully mixed solution was then recorded again. The procedure was repeated until 110 µl of HFiP were in total added to the peptide solution. The positions of the NH and the $C^{\alpha}H$ were used to monitor the association and conformational variation.

E.1.10 Diffusion experiments

Measurements were performed on the diastereoisomers of Z-Ala-Phe-Val-OMe using adequate concentrations as were used for the determination of the intramolecular association. Infrared spectroscopy was herein used for the concentration determination. The diffusion apparatus (Fig. E2) followed the method described by Anderson[268] and, thus, will be only described shortly: A capillary tube of a diameter of 1 mm and a volume of 0.4 ml was filled with the peptide solution using a syringe and placed vertically in a 50 ml vessel containing the pure solvent. The capillary was left for defined times in the solvent reservoir and was then removed and the concentration was determined IR spectroscopically. 15 hours were chosen as the duration of the experiments in carbon tetrachloride. After 60 min. the peptide concentration in the capillary was significantly decreased. Special care was taken to keep the temperature of the diffusion vessel and the solutions constant at 25 °C (±0.5 K). Preliminary tests have shown that in the time intervals chosen for the experiments the decrease of the concentration is linearly. Hence the diffusion rate v can be calculated from $\Delta c/\Delta t$ (Table A15).

The integrals of the free NH and NH_{C5} stretching vibration were used as a measure of the peptide concentration for toluene and carbon tetrachloride solutions, respectively. In carbon tetrachloride the CH stretching vibrations of the compounds are accessible and can be used as internal standard for the concentrations.

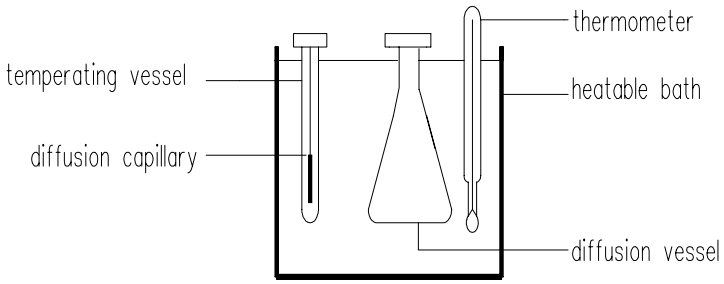


Fig. E2: Diffusion apparatus by Anderson

E.2 Additional data

In this part additional data are summarised which were used to create the plots, to receive equilibrium constants or fragmental values. To differentiate these tables and plots from previous data they were indicated by the letter „A“. The tables are ordered according to their occurrences in the text.

Table A1: $LogP_{Oct}$ values determined experimentally and calculated by the Rekker method for ACV and DCV derivatives[44]

Compound	logP (exp)	logP (calc)
ACV	-1.56	-1.55
NAcACV	-1.30	-1.26
OAcACV	-1.08	-0.58
diAcACV	-0.83	-0.29
DCV	-1.08	-1.05
OAcDCV	-0.61	-0.08

Table A2: Linear regression parameters for the CHI determination based on the φ_0 values of the standard set

Chromatographic column	r^2	Slope A	Intercept B
IAM	0.9958	27.626	-66.585
CD	0.9934	37.878	-104.34
CN	0.9951	23.012	-19.164
Pro	0.9843	30.454	-66.992

Table A3: CHI values obtained on various stationary phases[a]

Compounds	CHI_{IAM}	CHI_{CD}	CHI_{CN}	CHI_{In}	CHI_{Pro}
Paracetamol	**1.4**	**15.7**	**-2.5**	**10.1**	**20.8**
Acetophenone	**18.3**	**44.2**	**15.7**	**53.0**	**55.8**
Propiophenone	**26.6**	**51.8**	**23.6**	**65.4**	**67.1**
Butyrophenone	**32.2**	**55.7**	**31.3**	**74.8**	**75.9**
Valerophenone	**37.1**	**59.8**	**37.6**	**83.4**	**83.6**
Hexanophenone	**41.4**	**64.5**	**42.0**	**91.4**	**91.3**
Heptanophenone	**45.4**	**68.9**	**45.2**	**98.8**	**99.9**
Octanophenone	**49.2**	**73.2**	**47.4**	**107.0**	**109.4**
4-Iodophenone	45.5	70.8	21.3	60.4	60.7
Dibenzothiophene	55.4	77.9	46.2	101.5	101.6
4-Chlorophenol	38.9	59.5	9.2	53.1	54.6
4-Cyanophenol	26.6	45.1	0.1	34.1	41.7
Benzamide	1.9	24.0	-1.5	23.8	28.2
Caffeine	-4.5	7.0	8.2	20.1	22.2
Indasole	24.1	49.4	13.5	43.3	42.4
Anisole	27.2	56.0	8.8	67.1	67.8
Benzonitrile	19.6	47.6	4.5	55.5	56.5
Chlorobenzene	37.0	67.8	15.6	79.5	79.6
Naphthalene	42.8	68.0	36.2	86.0	84.5
Dinitrobenzene	23.8	54.2	4.6	58.8	58.7
Phenol	21.7	47.6	-4.3	38.2	43.5
Trifluoromethylphenol	38.4	56.1	18.4	58.6	59.0
Toluene	34.1	64.0	9.4	79.2	80.0
Corticosterone	30.9	39.6	32.7	53.5	50.2
Aniline	7.6	33.5	-4.7	38.4	43.2
Testosterone	39.4	58.4	36.7	69.6	64.6
Hydrocortisone-21-acetate	38.7	43.8	34.6	55.6	54.0
p-Toluidine	19.5	43.9	1.8	49.2	50.9
m-Nitroaniline	25.2	48.6	4.1	46.2	49.8

[a]In some cases CHI values are higher than 100 % which arises when compounds having longer retention than $\log k_0=0$ with 100 % organic phase and sometimes they have negative values which arises when compounds have a shorter retention than $\log k_0=0$ with 0% organic phase.

Table A4: Retention times t_R at different gradient times t_G, They were used for the CHI determination (see table 8) [in min.]

t_R at t_G [min]	2.5	3.5	5.5	7.5	10	15	30	45	60	90
Ac-Ala-OMe	3.635	3.854	4.212	4.501	4.834	5.303	6.402	7.307	7.856	9.222
Ac-D-Ala-OMe	3.632	3.847	4.211	4.507	4.829	5.314	6.423	7.302	7.891	9.178
Z-Ala-Ala-OMe	4.270	4.823	5.841	6.774	7.864	9.828	14.947	19.523	23.583	30.922
Z-Ala-Ala-Ala-OMe	4.094	4.611	5.558	6.424	7.446	9.277	14.116	18.469	22.394	29.370
Ac-Phe-OMe	4.147	4.652	5.563	6.392	7.348	9.033	13.356	17.151	20.466	26.878
Z-Phe-OMe	4.104	4.584	5.460	6.234	7.089	8.622	12.469	15.695	18.976	24.442
Z-Phe-Phe-OMe	5.048	5.821	7.316	8.760	10.504	13.768	22.750	31.049	39.110	54.384
Z-Phe-Phe-Phe-OMe	5.150	5.950	7.506	9.023	10.885	14.380	24.070	33.124	41.877	58.609
Ac-Val-OMe	3.880	4.288	4.983	5.580	6.247	7.376	10.157	12.468	14.511	18.075
Z-Val-Val-OMe	4.846	5.551	6.903	8.176	9.696	12.516	20.128	27.107	33.742	46.162
Z-Val-Val-Val-OMe	4.824	5.532	6.893	8.180	9.719	12.606	20.416	27.646	34.549	47.489
Z-Val-D-Val-Val-OMe	4.914	5.635	7.023	8.341	9.923	12.873	20.912	28.353	35.443	48.811
$100/t_G$	40	28.57	18.18	13.33	10.00	6.67	3.33	2.22	1.67	1.11
$1/100/t_G$	0.025	0.035	0.055	0.075	0.1	0.15	0.3	0.45	0.6	0.9

Table A5: Retention times t_R and bandwidth w for the phenylalanyl derivatives in dependence of t_G taken for the 230 nm reference band [in min.]

	Ac-Phe-OMe		Z-Phe-OMe		Z-Phe-Phe-OMe		Z-Phe-Phe-Phe-OMe	
t_G	t_R	w	t_R	w	t_R	w	t_R	w
2.5	4.147	0.0501	4.104	0.0729	5.046	0.0991	5.15	0.052
3.5	4.653	0.0394	4.584	0.0455	5.821	0.0692	5.95	0.0543
5.5	5.562	0.0443	5.459	0.0595	7.316	0.0891	7.505	0.062
7.5	6.391	0.047	6.234	0.0621	8.759	0.1025	9.025	0.0691
10	7.344	0.0537	7.089	0.0924	10.501	0.0508	10.887	0.0635
15	9.033	0.0556	8.622	0.0591	13.766	0.0589	14.375	0.0584
30	13.355	0.0744	12.47	0.0906	22.747	0.0858	24.063	0.0848
45	15.692	0.1456	15.692	0.1456	31.053	0.1065	33.119	0.1063
60	20.46	0.1107	18.993	0.1787	39.039	0.1321	41.897	0.1364
90	26.9	0.1462	24.422	0.2143	54.369	0.1756	58.578	0.1717

Table A6: Effect of the organic modifier on the retention time t_R measured on a ODS column Luna2 [in min.]

Substance		organic modifier		
	φ_0	AcN	MeOH	TFE
Ac-Ala-OMe	28.03	1.2235	1.401	1.444
Ac-D-Ala-OMe	27.77	1.2265	1.407	1.45075
Z-Ala-Ala-OMe	60.39	2.001	2.444	2.3795
Z-Ala-Ala-Ala-OMe	55.79	1.917	2.436	2.38835
Ac-Phe-OMe	56.98	1.8705	2.329	2.976
Z-Phe-OMe	56.62	1.822	2.2865	2.393
Z-Phe-Phe-OMe	78.36	2.644	2.974	3.179
Z-Phe-Phe-Phe-OMe	80.34	2.744	3.0805	3.4515
Ac-Val-OMe	47.33	1.591	1.999	2
Z-Val-Val-OMe	73.96	2.444	2.823	2.977
Z-Val-Val-Val-OMe	73.55	2.466	2.9085	3.02
Z-Val-D-Val-Val-OMe	75.36	2.5005	2.946	3.076

Table A7: Additional CHI values for peptide derivatives used for the calculation of descriptors

Compounds	CHI_{CD}	CHI_{IAM}	CHI_{CN}	CHI_{Pro}	CHI_{In}
Z-Ala-D-Trp-Val-OMe	45.78	29.01	44.52	64.84	61.80
Z-Ala-D-Phe-Val-OMe	42.65	26.10	43.44	65.63	65.54
Z-Ala-D-Met-Val-OMe	37.55	23.20	40.12	62.21	58.81
Z-Ala-D-Leu-Val-OMe	38.94	24.44	42.16	67.95	63.99
Z-Ala-D-Tyr-Val-OMe	38.12	24.32	39.21	54.81	50.99
Z-Ala-D-Val-Val-OMe	36.53	22.72	39.47	62.86	72.63
Z-Ala-D-His-Val-OMe	74.48	30.65	34.91	60.50	62.22
Z-Ala-D-Glu-Val-OMe	28.51	39.90	32.11	48.06	48.09
Z-Ala-D-Ser-Val-OMe	26.68	18.47	31.39	45.88	41.44
Z-Ala-D-Asn-Val-OMe	23.89	17.18	30.87	42.05	37.58
Z-Ala-D-Tyr(OB)-Val-OMe	51.29	32.60	49.23	78.86	77.45
Z-Ala-D-Lys(NZ)-Val-OMe	47.28	29.97	46.50	73.89	72.07
Z-Ala-D-Asp(OB)-Val-OMe	44.83	27.91	45.61	71.20	69.22
Z-Ala-D-Ser(OB)-Val-OMe	43.54	27.85	44.98	70.64	68.44
Z-Ala-D-Arg(Tos)-Val-OMe	42.58	29.48	43.30	62.01	61.94
Z-Ala-L-Val-Gly-OEt	31.53	17.72	-	53.30	48.00
Z-D-Ala-Gly-Val-OMe	30.36	19.06	-	51.83	45.54
Z-Ala-D-Val-Gly-OEt	32.04	18.26	33.02	52.51	48.56
Z-Ala-D-Val-Ala-OMe	30.91	17.68	-	53.92	48.25
Z-Gly-Ala-Gly-OMe	23.65	8.71	19.38	41.67	41.53
Z-Gly-L-Ala-Val-OMe	30.41	17.93	31.77	-	44.69
Z-Gly-L-Ala-Phe-OMe	37.88	24.33	38.95	55.92	52.08
Z-Gly-D-Ala-Phe-OMe	37.44	23.81	39.21	55.88	51.66
Z-Gly-L-Phe-Gly-OEt	38.76	24.25	-	57.48	52.63
Z-Gly-D-Phe-Gly-OEt	38.73	24.00	38.80	55.99	52.45
Z-Ala-Val-Tyr-OMe	34.29	26.34	39.31	57.82	57.75
Z-Ala-Val-Pro-OMe	35.98	13.50	35.43	58.69	-
Z-Ala-D-Val-Leu-OMe	39.09	25.22	42.16	68.04	64.20
Z-Leu-Val-Val-OMe	38.24	23.08	42.80	74.40	71.12
Z-Ala-Val-Val-OBzl	47.34	29.50	47.00	74.74	72.63
Z-Ala-D-Val-Val-OBzl	47.36	30.49	47.24	76.94	72.63
Z-Ala-Val-Val-OBu	42.84	27.35	44.51	75.58	73.32
Z-Ala-Phe-Val-OBu	47.34	30.79	48.61	81.32	79.29
Boc-Ala-Phe-Val-OMe	36.53	22.33	39.54	-	62.36
Boc-Ala-Pro-Val-OMe	24.11	8.50	27.40	55.08	54.98

(Table 7, continued)

Z-Ala-Val-OMe	36.51	21.08	33.38	58.92	55.47
Z-Ala-Phe-OMe	47.27	29.15	41.34	57.54	67.29
Z-Phe-Val-OMe	51.56	32.19	46.00	75.24	75.01
Z-Ala-Ser-OMe	58.31	9.87	47.92	42.52	42.33
Z-Ala-Trp-OMe	51.24	32.50	43.09	65.64	65.56

Table A8: Additional experimental descriptors received from the CHI values, The descriptors of the tripeptide derivatives were used for the calculation of fragmental descriptors

Compound	R_2	π_2^H	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x
Z-Ala-L-Val-Gly-OEt	1.138	4.78	0.29	1.77	3.169
Z-Ala-D-Val-Gly-OEt	1.138	4.48	0.35	1.81	3.169
Z-Ala-D-Val-Ala-OMe	1.123	4.70	0.28	1.79	3.169
Z-Gly-Ala-Gly-OMe	1.195	3.72	0.08	1.68	2.605
Z-Gly-L-Ala-Val-OMe	1.141	4.03	0.50	1.80	3.028
Z-Gly-D-Ala-Phe-OMe	1.856	5.09	0.31	1.88	3.354
Z-Gly-L-Ala-Phe-OMe	1.856	5.11	0.32	1.86	3.354
Z-Gly-L-Phe-Gly-OEt	1.872	5.51	0.25	1.79	3.354
Z-Gly-D-Phe-Gly-OEt	1.872	5.19	0.31	1.85	3.354
Z-Ala-Val-Phe-OMe	1.807	6.18	0.16	1.95	3.776
Z-Ala-Val-Leu-OMe	1.096	5.31	0.21	1.87	3.591
Z-Ala-Val-Tyr-OMe	2.047	6.16	0.34	2.12	3.835
Z-Ala-Val-Pro-OMe	1.482	5.30	0.00	1.84	3.342
Z-Ala-D-Val-Leu-OMe	1.096	5.25	0.20	1.86	3.591
Z-Leu-Val-Val-OMe	1.064	5.95	0.00	1.97	3.873
Z-Ala-Val-Val-OBzl	1.766	6.91	0.03	1.95	4.058
Z-Ala-D-Val-Val-OBzl	1.766	6.91	0.02	1.94	4.058
Z-Ala-Val-Val-OBu	1.083	6.03	0.07	1.89	3.873
Z-Ala-Phe-Val-OBu	1.768	7.09	0.00	1.98	4.199
Boc-Ala-Phe-Val-OMe	1.096	5.32	0.18	1.90	3.591
Boc-Ala-Pro-Val-OMe	0.792	4.37	0.00	1.89	3.157
Z-Ala-Val-OMe	0.969	3.15	0.23	1.50	2.630
Z-Ala-Phe-OMe	1.670	4.35	0.13	1.50	2.956
Z-Phe-Val-OMe	1.640	5.00	0.08	1.54	3.238
Z-Ala-Ser-OMe	1.215	3.62	0.32	1.25	2.407
Z-Ala-Trp-OMe	2.549	5.42	0.14	1.60	3.186

Table A9: Dipeptide descriptors predicted from fragmental descriptors

Compound	R_2	π_2^H	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x
Z-Ala-Ala-OMe	1.042	3.62	0.12	1.17	2.313
Z-Val-Val-OMe	0.986	4.91	0.09	1.24	2.877
Z-Phe-Phe-OMe	2.426	7.11	0.03	1.44	3.529
Z-Ala-Phe-OMe	1.734	5.37	0.08	1.31	2.921
Z-Phe-Val-OMe	1.706	6.01	0.06	1.34	3.203
Z-Ala-Ser-OMe	1.263	3.63	0.45	1.34	2.372
Z-Ala-Trp-OMe	2.587	6.17	0.08	1.43	3.151

Table A10: CHI values of the ureido sugar derivatives used for the calculation of the descriptors

Compound	CHI_{IAM}	CHI_{CD}	CHI_{CN}	CHI_{In}	CHI_{Pro}
-Gly-OEt	-9.26	24.82	17.37	49.28	65.36
-Ala-OEt	2.21	35.18	24.11	52.88	67.36
-L-Val-OEt	11.15	42.88	33.49	60.98	72.94
-D-Val-OEt	23.38	27.23	21.96	53.83	67.93
-Leu-OEt	17.20	48.51	37.02	66.15	75.32
-Ile-OEt	17.43	48.78	37.14	66.04	75.30
-Phe-OMe	18.44	51.95	39.03	66.01	73.59
-Gly-Ala-OEt	-8.45	22.60	33.65	48.55	64.80
-Ala-Gly-OEt	-15.89	20.64	20.29	47.45	64.06
-Ala-Ala-OEt	-6.64	24.35	25.00	50.37	65.86
-Gly-Val-OEt	10.67	34.99	33.15	55.76	69.31
-Val-Gly-OBzl	3.51	30.51	31.08	52.65	67.60
-Leu-Gly-OEt	9.76	36.27	33.30	56.77	69.86
-Gly-Phe-OBzl	16.65	44.54	37.25	57.28	70.40
-Phe-Gly-OEt	16.40	46.48	37.55	59.24	71.48
-Ala-Phe-OBzl	28.94	64.03	45.29	71.03	78.73
-Gly-Val-OBzl	16.69	44.50	36.73	57.17	70.39

Table A11: Descriptors of the ureido sugars- experimentally derived and calculated

	experimentally determined descriptors					predicted descriptors				
Compound	R_2	π_2^H	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x	R_2	π_2^H	$\sum\alpha_2^H$	$\sum\beta_2^H$	V_x
-Gly-OEt	0.871	5.56	0.00	1.72	3.167	0.935	6.11	0.20	1.53	3.150
-Ala-OEt	0.854	5.99	0.00	1.64	3.307	0.912	6.34	0.12	1.56	3.291
-L-Val-OEt	0.846	6.64	0.00	1.66	3.589	0.884	6.99	0.10	1.60	3.572
-D-Val-OEt	0.846	6.05	0.12	1.82	3.589	0.889	6.94	0.12	1.61	3.572
-Leu-OEt	0.859	6.78	0.00	1.65	3.730	0.889	7.16	0.13	1.65	3.713
-Ile-OEt	0.859	6.79	0.00	1.64	3.730	0.908	6.76	0.08	1.65	3.787
-Phe-OMe	1.624	7.23	0.00	1.67	3.774	1.611	7.74	0.03	1.70	3.758
-Gly-Ala-OEt	1.114	5.82	0.00	2.18	3.705	1.038	6.34	0.24	2.11	3.723
-Ala-Gly-OEt	1.114	6.73	0.00	2.06	3.705	1.038	6.34	0.24	2.11	3.723
-Ala-Ala-OEt	1.098	7.15	0.00	1.99	3.846	1.016	6.58	0.16	2.15	3.864
-Gly-Val-OEt	1.1	7.16	0.00	1.99	3.987	1.010	6.99	0.23	2.15	4.005
-Val-Gly-OBzl	1.886	8.68	0.00	2.28	4.453	1.691	8.42	0.15	2.25	4.472
-Leu-Gly-OEt	1.099	7.61	0.00	2.02	4.127	1.015	7.16	0.25	2.20	4.146
-Gly-Phe-OBzl	2.694	10.16	0.00	2.23	4.779	2.411	9.52	0.12	2.35	4.798
-Phe-Gly-OEt	1.906	8.69	0.00	1.98	4.313	1.730	8.09	0.20	2.25	4.331
-Ala-Phe-OBzl	2.688	10.96	0.00	1.96	4.920	2.388	9.76	0.04	2.39	4.939
-Gly-Val-OBzl	1.886	9.03	0.00	2.07	4.453	1.691	8.42	0.15	2.25	4.472

Table A10: Hydrogen bond acidity and basicity of proton bases and acids and the equilibrium constant for the association in methylene chloride[166, 269, 270]. (First column indicates the grouping of the compounds)

	Acid	Base	α_2^H	β_2^H	$\alpha_2^H*\beta_2^H$	logK
A	p-Fluorophenol	Tetrahydrothiophene	0.629	0.2644	0.1663	-0.15
A	p-Fluorophenol	Dioxane	0.629	0.475	0.2988	0.51
A	p-Fluorophenol	THF	0.629	0.51	0.3208	0.75
A	HFP	THF	0.771	0.51	0.3932	1.61
B	p-Fluorophenol	4-Nitroanisol	0.629	0.345	0.2170	-0.39
B	p-Fluorophenol	Diethylcarbonate	0.629	0.45	0.2830	0.22
B	p-Fluorophenol	N.N dimethylthioacetamide	0.629	0.4922	0.3096	0.48
B	p-Fluorophenol	Anthrone	0.629	0.506	0.3183	0.37
B	p-Fluorophenol	Diphenylsulfone	0.629	0.512	0.3220	0.22
B	p-Fluorophenol	N.N-dimethylmethanesulfonamide	0.629	0.517	0.3252	0.32
B	p-Fluorophenol	Cyclohexanone	0.629	0.523	0.3290	0.59
B	p-Fluorophenol	Xanthone	0.629	0.531	0.3340	0.6
B	p-Fluorophenol	Triethylphosphine sulfide	0.629	0.552	0.3472	0.63
B	p-Fluorophenol	DMF	0.629	0.6628	0.4169	1.3
B	p-Fluorophenol	Formamide	0.629	0.6628	0.4169	0.97
B	p-Fluorophenol	DMF	0.629	0.663	0.4170	1.18
B	p-Fluorophenol	Diphenyl sulfoxid	0.629	0.666	0.4189	1.17
B	p-Fluorophenol	Benzamide	0.629	0.674	0.4239	1.02
B	1-Naphthol	Tetramethylurea	0.608	0.743	0.4517	1.41
B	p-Fluorophenol	Acetamide	0.629	0.7297	0.4590	1.24
B	p-Fluorophenol	Trimethylphosphate	0.629	0.7616	0.4790	1.42
B	p-Fluorophenol	DMSO	0.629	0.774	0.4868	1.57
B	p-Fluorophenol	Dibutyl sulfoxid	0.629	0.7851	0.4938	1.73
B	HFiP	N-Methylacetamide	0.771	0.72	0.5551	2.62
B	p-Fluorophenol	Triphenylphosphineoxide	0.629	0.919	0.5780	2.07
B	p-Fluorophenol	Hexamethylphosphotriamide	0.629	1	0.629	2.53
B	p-Fluorophenol	Triphenylarsine oxide	0.629	1.027	0.6460	2.94
C	TFPE	Z-Phe-OMe	0.733		0	0.12
C	TFPE	Z-Val-OMe	0.733	0.5206	0.3816	0.09
C	HFiP	Acetone	0.771	0.497	0.3832	1.20
C	HFiP	Z-Val-OMe	0.771		0	
C	HFiP	Ac-Ala-OMe	0.771	0.698	0.538	2.11
B	HFiP	N.N-Dimethylacetate	0.771	0.78	0.601	2.96
B	p-Fluorophenol	Trimethylphosphine oxide	0.629	0.9801	0.616	2.07

TFPE= 2-phenyl-1.1.1-trifluoroethanol

Table A13: Integral molar extinction coefficients of selected tripeptide derivatives dissolved in methylene chloride [in l/mol*mm]

Substance	ε_{Mol}			
	at -10 °C	at -5 °C	at 15 °C	at 25 °C
Z-Phe-Ala-Val-OMe	22.13	21.71	20.3	19.58
Z-Ala-Phe-Val-OMe				
LLL	22.66	22.36	21.00	20.39
LDL	22.15	22.04	21.04	20.11
DLL	22.13	22.02	20.9	20.52
Z-Ala-Phe-Val-OtBu	22.21	21.92	20.48	20.08
Z-Ala-Leu-Val-OMe				
LLL	23.61	23.29	22.26	21.41
LDL	23.23	23.38	21.77	21.09

Table A14: Normalised integral molar extinction coefficients of selected tripeptide derivatives dissolved in carbon tetrachloride used for the approximation of the intramolecular equilibrium constant K [in l/mol*mm]

Substance	ε_{Norm}				
	at -10 °C	at -5 °C	at 15 °C	at 25 °C	at 40 °C
Z-Phe-Ala-Val-OMe	18.10	18.45	18.68	18.45	17.91
Z-Ala-Phe-Val-OMe					
LLL	6.63	7.01	10.43	12.21	12.57
LDL	7.41	7.99	10.23	10.91	11.59
DLL	6.70	6.74	10.41	12.05	13.24
Z-Ala-Phe-Val-OtBu	7.22	7.70	10.43	12.21	12.57
Z-Ala-Leu-Val-OMe					
LLL	5.86	6.17	9.74	11.37	12.86
LDL	6.97	7.63	9.43	10.03	10.60

[a] ε_{Norm} was received by dividing the integral absorbance by d*c. If no intramolecular association occurs like in Z-Phe-Ala-Val-OMe similiar values would be expected also for the other compounds. The loss of ε_{Norm} represents the concentration of the associated NH functions.

Table A15: Differences in the peptide concentration Δc after diffusion measured on the intensity decrease of the NH_{C5} and CH_2 signals [in 10^{-5} mol/l]

	Carbon tetrachloride		Toluene
Substance	NH_{C5}: Δc (after 15 h)	$CH_{2\ as}$: Δc (after 15 h)	NH: Δc (after 60 min.)
Z-Phe-Ala-Val-OMe	20.81	18.9	-
Z-Ala-Phe-Val-OMe			
LLL	8.96	7.29	46.79
LDL	15.94	17.76	57.76
DLL	7.77	6.8	42.21
Z-Ala-Leu-Val-OMe			
LLL	17.66	17.87	123.43
LDL	14.68	16.63	79.0
DLL	15.97	12.09	79.7

Table A16: Hydrogen bond basicity scales β_2^H and $\sum\beta_2^H$ received from different methods

Substance	$\sum\beta_2^H$ (HPLC)	$\sum\beta_2^H$ (peptide fragment)	$\sum\beta_2^H$ (UCL)	β_2^H(OH)	β_2^H(sum)
Ac-Ala-OMe	1.10	0.59	1.04	0.69	0.64
Ac-Phe-OMe	0.62	0.72	1.09	0.69	0.64
Ac-Val-OMe	0.91	0.62	1.06	0.69	0.64
Z-Ala-Ala-OMe	0.90	1.18	1.65	0.69	0.52
Z-Phe-Phe-OMe	1.16	1.44	1.75	0.68	0.53
Z-Val-Val-OMe	0.97	1.24	1.69	0.69	0.55
Z-Ala-Val-OMe	1.50	1.21		0.68	0.57
Z-Ala-Ala-Ala-OMe	1.28	1.77	2.26	0.78	0.60
Z-Phe-Phe-Phe-OMe	1.61	2.16	2.4	0.72	0.60
Z-Val-Val-Val-OMe	1.33	1.86	2.31	0.72	0.60
Z-Ala-Val-Val-OMe	1.87	1.83		0.75	0.60
with Phenol					
Ac-Val-OMe	0.91	0.62	1.06	0.77	0.71
Z-Val-Val-OMe	0.97	1.24	1.69	0.72	0.50
Z-Val-Val-Val-OMe	1.33	1.86	2.31	0.74	0.69

Erklärung

Ich erkläre, daß ich die vorliegende Arbeit selbständig und ohne fremde Hilfe angefertigt habe. Es wurden keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen sind als solche kenntlich gemacht.

Diese Habilitationsschrift wurde bisher an keiner anderen Universität oder Hochschule vorgelegt.

Halle (Saale), Juni 2000

Dr. M. Plass

Diese Arbeit wurde von April 1995 bis Juni 2000 am Institut für Physikalische Chemie der Martin-Luther-Universität Halle-Wittenberg angefertigt. Teilergebnisse, die während der Forschungsaufenthalte am University College London erzielt wurden, sind in die Arbeit eingeflossen.

Meinem akademischem Lehrer, Herrn **Prof. Dr. A. Kolbe**, danke ich ganz besonders herzlich für die Anregungen und zahlreichen Diskussionen.

Für ihre stets zuverlässige und umsichtige Arbeit und Hilfe im Laboratorium möchte ich Frau **I. Schaller** danken.

Die Diffusionsuntersuchungen an Peptidderivaten wurden von Frau **Dr. A. Gröbel** im Rahmen ihrer Dissertation durchgeführt. Ihr und den ausländischen IAESTE Studenten, die in unserem Laboratorium ein Praktikum absolviert haben, möchte ich für die engagierte Arbeit danken.

Die in dieser Arbeit verwendeten Di- und Tripeptide wurden von Frau **Prof. Dr. C. Griehl**, Fachhochschule Köthen, und ihren Mitarbeitern synthetisiert und mir zur Verfügung gestellt. Für die gute Zusammenarbeit in den letzten Jahren möchte ich ganz herzlich danken.

Frau **Prof. Dr. I Wawer, Frau Dr. M. Weychert**, Medizinische Fakultät Warschau, und Herrn **Dr. D. Ströhl** möchte ich für die NMR Analysen danken.

Frau **Prof. Dr. T. B. Freedman** und Herrn **Prof. Dr. L. A. Nafie**, University Syracuse, danke ich für die VCD Untersuchungen.

Herrn **Prof. Dr. A. Kristl**, Universität Ljubljana, danke ich für die Bereitstellung der Acyclovirderivate.

Herrn **PD Dr. R. Friedemann** danke ich für die Bereitstellung von Literatur zu den quantenmechanischen Berechnungen an Diolen.

Mein herzlicher Dank gilt auch Herrn **Prof. Dr. M. H. Abraham**, University College London, für die freundliche Aufnahme in seiner Arbeitsgruppe während meiner Aufenthalte in London und für die hilfreichen Diskussionen der QSAR Parameter.

Frau **Prof. Dr. K. Valko** und Herrn **Prof. Dr. D. Reynolds** gaben mir die Möglichkeit, wichtige Untersuchungen in den Laboratorien der Physical Science Unit des GlaxoWellcome Medicine Research Centre, Stevenage, UK, durchzuführen. Ihnen und allen Mitarbeitern gilt mein herzlicher Dank.

Für die großzügige finanzielle Unterstützung der Arbeiten danke ich dem **Fond der Chemischen Industrie**, der **Deutschen Forschungsgemeinschaft**, der **Royal Society of Chemistry** sowie der **Deutschen Akademie der Naturforscher Leopoldina**.

CURRICULUM VITAE

Name	Monika Plass, Dr. rer. nat.
geboren am	3. Juni 1965 in Potsdam

Schulbildung

1972-1980	Polytechnische Oberschule in Bergholz-Rehbrücke
1980-1984	Erweiterte Oberschule „Alexander und Wilhelm Humboldt“ in Potsdam

Studium

1985-1990	Chemiestudium an der Martin-Luther-Universität Halle-Wittenberg
	Diplom bei A. Kolbe

Beruflicher Werdegang

1990-1993	Aspirantin am Institut für Analytische Chemie
	der Martin-Luther-Universität Halle-Wittenberg
	Promotion bei A. Kolbe
1993-1995	Postdoktorandin an der University of North Texas, Denton, TX, USA
	bei W. K. Brostow
seit 1995	Habilitandin am Institut für Physikalische Chemie
	der Martin-Luther-Universität
1996 (6 Monate)	Forschungsaufenthalt am University College London
1999 (2 Monate)	bei M. H. Abraham
1993-1995	DFG-Stipendium
1994	Christian-Wolff-Preis der Martin-Luther-Universität
1996	Leopoldina-Förderpreis der Deutschen Akademie der Naturforscher
1999	„Grant for International Authors“ of the Royal Society of Chemistry