STRUCTURE AND CONFORMATIONAL BEHAVIOUR OF PEPTOID PEPTIDOMIMETICS

Towards an Understanding of the Mechanism of Action of Substance P and NK_1 Antagonists

Structuur en conformationeel gedrag van peptoïde peptidomimetica
Naar begrip van het werkingsmechanisme van substance P en NK_1 antagonisten

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de Rector Magnificus, Prof. dr H. O. Voorma
ingevolge het besluit van het College van Decanen
in het openbaar te verdedigen
op 5 september 1997 des middags te 12.45 uur

door

Gerrit Jan Boks

geboren op 21 augustus 1969, te Elst (Gld.)

promotores: Prof. dr J. P. A. E. Tollenaere

verbonden aan het Utrecht Institute for Pharmaceutical Sciences
van de faculteit Farmacie van de Universiteit Utrecht.

Prof. dr J. Kroon

verbonden aan het Bijvoet Centrum van de faculteit Scheikunde
van de Universiteit Utrecht.

Dit proefschrift werd mede mogelijk gemaakt met financiële steun van

Janssen-Cilag B.V., Tilburg
Dr Saal van Zwanenbergstichting, Oss

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Boks, Gerrit Jan

Structure and conformational behaviour of peptoid peptidomimetics. Towards an understanding of the mechanism of action of substance P and NK_1 antagonists / Gerrit Jan Boks.
- Utrecht: Universiteit Utrecht, Faculteit Farmacie.
Proefschrift Universiteit Utrecht.
- Met lit. opg. - Met samenvatting in het Nederlands.
ISBN 90-393-1661-9
Trefw.: peptidomimetica / peptoïden / substance P / NK_1 antagonisten

List of Abbreviations

Amino-acid residues and derivatives

Ala	alanine
Arg	arginine
Asn	asparagine
Asp	aspartic acid
Ava	δ-amino valeric acid
Cys	cysteine
Gln	glutamine
Glu	glutamic acid
Gly	glycine
His	histidine
Ile	isoleucine
Leu	leucine
Lys	lysine
Met	methionine
pGlu	pyroglutamic acid
Phe	phenylalanine
Pro	proline
Sar	sarcosine, *N*-methylglycine
Ser	serine
Thr	threonine
Trp	tryptophan
Tyr	tyrosine
Val	valine

Miscellaneous

+ve	positive potential extremum
-ve	negative potential extremum
2D	two-dimensional
3D	three-dimensional
6-31G*	Gaussian basis set for the description of atomic orbitals in molecular orbital calculations.
Ac	acetyl, or acceptor group in combination with a donor.
ACE	angiotensin-converting enzyme
C_1	molecular point group with no symmetry at all.
CCK-B	cholecystokinin-B
CFF91	consortium force field 1991
C_S	molecular point group with only mirror-image symmetry
CSD	Cambridge Structural Database
CVFF	consistent valence force field
D	donor
D-AcSP	*N*-acetyl substance P consisting of (D)-amino acids only
DMSO-d_6	deuterated dimethylsulfoxide
FAB	fast atom bombardment
Fmoc	9-fluorenylmethyloxy carbonyl
GED	gas electron-diffraction
HF	Hartree-Fock
HIV-1	human immunodeficiency virus type 1
HPLC	high pressure/performance liquid chromatography
IC_{50}	here: concentration necessary to inhibit 50% of the maximal agonist binding.
IUB	International Union of Biochemistry
IUPAC	International Union of Pure and Applied Chemistry
K_i	equilibrium dissociation constant for an antagonist or an inhibitor
L-AcSP	*N*-acetyl substance P consisting of (L)-amino acids only
$[M+H]^+$	molecular peak position in mass spectrometry, including an additional proton.
m/z	molecular mass over charge ratio in mass spectrometry
MEP	molecular electrostatic potential
MP2	Møller-Plesset perturbation theory with 2nd. order correction.
MPD	2-methyl-2,4-pentanediol
N^+	unitary positively charged nitrogen atom
NAla	*N*-methylglycine, peptoid analogue of alanine
NEP	neutral endopeptidase

List of Abbreviations

NHMe	*N*-methylamino (-amide)
NK_1	neurokinin-1
NK_2	neurokinin-2
NK_3	neurokinin-3
NKA	neurokinin A
NKB	neurokinin B
NMe_2	dimethylamino (-amide)
NMR	nuclear magnetic resonance
NPγ	neuropeptide γ
NPK	neuropeptide K
NXxx	generic peptoid residue, analogous to amino acid Xxx
PDB	Brookhaven Protein Data Bank
PEG	polyethylene glycol
RMS	root-mean-square
RMSd	root-mean-square distance
RNA	ribonucleic acid
SAR	structure-activity relationship
SP, SP_{1-11}	substance P
SP_{6-11}	minimal sequence of substance P
TAR	transactivator-responsive element
TM	transmembrane domain
XED	extended electron distribution
Xxx	three-letter notation for an amino acid.

Geometrical descriptors

d(A-B)	length of the vector AB or interatomic distance of atoms A and B.
<(A-B-C)	angle between the vectors BA and BC. In the case of atoms: bond angle A-B-C.
τ(A-B-C-D)	dihedral angle of the planes through ABC and BCD. In the case of atoms: torsion angle.

Chapter 1

General Introduction

Macromolecular targets that depend on the recognition of peptides or peptide epitopes for their function are ubiquitous in nature. The peptide ligands associated with these processes include signalling peptides, such as the sensory neuropeptides (e.g. substance P),[1] peptide hormones (e.g. fertility hormones),[2] and also peptides that are released from cells of the immune system (e.g. interleukins, cytokines) which trigger other cells to combat infection.[3] The role of some of these peptides in pathophysiological processes in the human body (e.g. in pain transmission or inflammation) stresses the importance of their receptors as attractive possible targets for drug design; a blockade of the corresponding peptide receptors impairs these processes. On the other hand, stimulation of specific peptide hormone receptors, plays an important role e.g. in treating infertility.[2] Meanwhile, peptides also play an important role as degradation product of proteins. Exogenous proteins stemming from invading viruses, as well as endogenous proteins in both the infected and normal cells, are constantly degraded into peptides. These peptides are then, as part of the class II major histocompatibility complex at the cell membrane, presented to T-lymphocytes so that proper action of the immune system is elicited.[3] Under normal circumstances only the non-self peptides are recognized by T-cells, whereas in autoimmune diseases also some self-peptides activate the immune system and as a consequence also healthy cells are destroyed. Compounds that interfere with these recognition processes may be employed to study the underlying mechanisms and ultimately lead to a therapeutically useful compound. Other recognition processes are protein-protein interactions, such as involved in the action of many antibodies,[4] in which peptides are in fact constituting epitopes. For instance, upon infection with the Gram negative bacterium *Neisseria meningitidis*, which can cause life-threatening meningitis and sepsis in humans, also protective antibodies are formed against one or more of the extracellular loops of an outer membrane protein in the bacterium.[5] A compound that is able to mimic the action of this antigenic loop(s) may be used as a vaccine to provide immunity prior to bacterial infection.

In conclusion, a general strategy aimed at identifying so-called peptidomimetic compounds[6] that can interfere with these and other peptide recognition processes would be highly valuable. Biologically active compounds thus identified may not only be used as pharmacological tools in unravelling the pathophysiological processes, but may eventually

lead to compounds to be used in therapy. As a fundamental step in elucidating the mechanism of action of peptidomimetics, the structure and conformational behaviour of one class of peptidomimetics, *viz.* peptoids, in relation to the tachykinergic NK_1 receptor system, are addressed in this thesis.

Peptoid Peptidomimetics

Crucial to the processes mentioned above is the molecular recognition of a peptide by a macromolecular receptor. Peptides themselves seem to be the first candidates to interfere with this action. The use of peptides in drug therapy, however, usually suffers from a number of drawbacks.[7] In general, peptides have a poor oral-bioavailability due to degradation in the stomach and intestines, as well as a poor absorption in the intestinal system.[7] Also, peptides often have a short duration of action because they are rapidly degraded by proteolytic enzymes in the blood and in other tissues and fail to cross the blood-brain-barrier.[7] Compounds to be used in therapy, but also in research, are in the best case devoid of these disadvantages. Non-peptide peptidomimetics may play an important role in that respect.[6] Peptidomimetics are compounds that can mimic the critical features of the molecular recognition process of the parent peptide and thereby block or reproduce the action of the peptide. Peptidomimetics can in this definition thus be either antagonists or agonists. Probably the oldest example of a non-peptide peptidomimetic agonist for a peptide receptor system is morphine, that mimics the opioid peptides.[6] Meanwhile, an antagonist is in our opinion only considered a peptidomimetic if the antagonist and the peptide agonist bind at a similar site at the receptor and thus mimicry of the binding exists.

To make peptidomimetics a useful part of a general ligand-design strategy a number of aspects has to be taken into account. The synthesis of compounds that may perform the role of peptidomimetic should be relatively straightforward and amenable to automation.[8] Molecular diversity can be obtained by the combination of relatively simple molecular building blocks which can be combined to form any peptidomimetic (see e.g. ref. 9). The modular build-up of peptides, as well as the chemical functionality of the peptide that is necessary for exerting its action, is thus preferably maintained in peptidomimetics. As a consequence, with a limited number of building blocks a great diversity of compounds, a library of compounds, can be made. The amide bonds in peptides are very susceptible to enzymatic hydrolysis and consequently the peptide rapidly looses its activity. Yet, various alternative linking groups for the building blocks[9] can be employed, that are expected to be

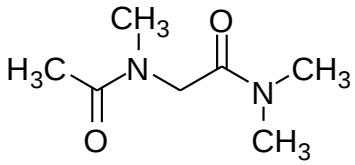
less susceptible to enzymatic hydrolysis while maintaining the side-chain functionality of the parent peptides.[9] The general rationale behind the design of these peptidomimetics is the assumption that the side-chain functional groups are mainly responsible for the specific interactions with the macromolecular receptor, and that the peptide linkage functions as a scaffold that can be replaced by another group without a severe loss of activity at the receptor. Here we focus on peptoids,[10,11] oligomers of *N*-substituted glycines, as one of the relatively new classes of peptidomimetics.

Peptoids at first sight resemble peptides to a considerable degree (Fig. 1). The topological positions of the side chains in peptides are maintained in peptoids; the same number of bonds connect the side-chain attachment points in the backbones of both peptoids and peptides.[10] This might enable peptoids to adopt the same three-dimensional arrangement of the side-chain functional groups as the parent peptide in its bioactive conformation, and thereby mimic the peptide's action. When the side-chain positions of peptides and peptoids have been topologically aligned (Fig. 1, **A** and **B**), the positions of the backbone carbonyl groups relative to the side chains positions, differ, however, in peptides and peptoids. Yet, the peptide sequence can also be translated from the C- towards the N-terminus (in addition to a direct N- to C-terminus translation) resulting in a *retro*peptoid (Fig. 1, **C**).[11] The topological similarity of the side chain positions is thereby conserved and, moreover, also the relative positions of the backbone carbonyl groups with respect to the side chains in the backbone is identical to the situation in the parent peptide (Fig. 1, **C**). Therefore, it is expected that retropeptoids may generally be the preferred peptoid for binding by a macromolecular receptor. In an attempt to standardize the communication with respect to the molecular structure of peptoids, tentative rules for describing their conformation were devised (see Appendix A). The established nomenclature for peptides[12] thereby served as a basis.

In peptoids the amino-acid side chains are on the amide nitrogen atoms instead of on the C_α carbon atom in peptides. In addition to the 18 peptoid monomers that can be 'constructed' based on the proteinogenic amino acids, proline and glycine remain unaltered. The nature of the side chains is, however, not restricted to the ones commonly encountered in peptides and proteins, but can e.g. be based on the structure of known ligands for a specific (non-peptide) receptor[13] or tailored to obtain a maximum degree of diversity in the functional groups present.[14] By attaching the side chain to the amide nitrogen both the chirality at the C_α carbon atom (except for proline) and the amide NH hydrogen bond donor group (except for glycine) are lost. Given the importance of the amide NH group in the

formation of hydrogen bond stabilized secondary structural elements in peptides or in peptide-protein interactions,[15] this loss may have unfavourable consequences. The loss of the NH hydrogen, on the other hand, also results in the (required) increased stability of peptoids towards peptide degrading enzymes.[10]

In the past, the structure and preferred conformations of amino acids, peptides, peptide epitopes, and proteins has been the subject of extensive research.[16,17] However, in contrast to the considerable knowledge on peptides, very little is known about the structure and conformational behaviour of peptoids. Based on empirically calculated Ramachandran-type plots[16] for peptoid monomers Simon et al. have suggested that the conformational behaviour of the peptoid backbone is largely unaffected by the nature of the side chains.[10] Also a greater diversity of conformational states for a peptoid compared to that of a peptide has been suggested. This effect is mainly due to the lack of a chiral carbon atom.[10] Furthermore, *cis/trans* isomerization of the tertiary amide bonds as present in peptoids has been observed in NMR spectra at room temperature in different solvents.[18] At the NMR time scale this implies that the barrier for rotation about the tertiary amide bonds is certainly higher than ca. 4.8 kcal/mol,[19] but lower than the barrier for rotation around a secondary amide bond. Recently, molecular dynamics studies as well as a quantum chemical study on the local minimum energy conformations of a capped model peptoid monomer, Ac-NAla-NMe_2 (see formula) *in vacuo*, have been reported.[20,21] In addition, as far as experimental structural studies on peptoids are concerned, we are mainly restricted to studies on peptides that also contain peptoid residues. Structural chemical knowledge on peptoid *oligomers* at atomic resolution is absent, however.



Recently, a number of biologically active peptoids has been reported. Screening of a library of ca. 5000 di- and tripeptoids has resulted in ligands with nanomolar affinities (K_i) for the α_1-adrenergic and μ-opiate receptors.[13] In this case, the peptoid side chains had been biased towards the structures of known ligands for a number of G-protein coupled receptors.[13] Note in this respect that the endogenous ligands for the α_1-adrenergic receptor

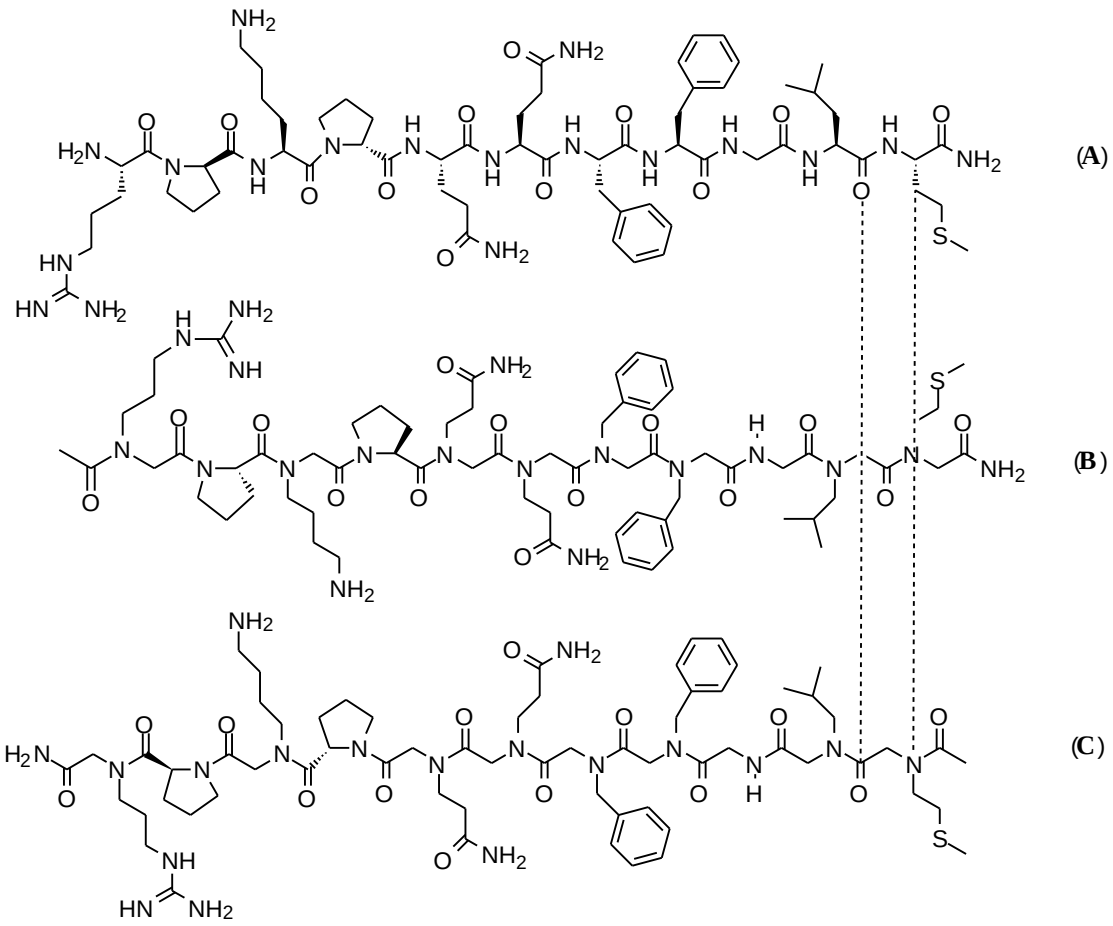


FIGURE 1. Structural formula of the undecapeptide substance P (SP, **A**), its peptoid (**B**) and its retropeptoid (**C**).

are catecholamines, rather than peptides. Previously, Simon et al., reported on peptoid inhibitors for bovine pancreatic α-amylase and the hepatitis A viral 3C protease.[10] Furthermore, an octapeptoid completely consisting of the non-proteinogenic residue *N*-(aminoethyl)glycine, as a mimic of lysine residues, was found to bind to the transactivator-responsive element (TAR) RNA of HIV-1.[10] This peptide was derived from an epitope of the corresponding TAR RNA binding protein.[10] In the case of α-amylase, the peptoid and retropeptoid sequences (with sometimes higher homologues of amino-acid side chains) were based on a tripeptide epitope of the protein inhibitor tendamistat and in the case of the hepatitis A protease, on the octapeptide consensus sequence of substrates. The retropeptoid inhibitors for both the glycolytic enzyme α-amylase and the hepatitis A virus 3C protease revealed a higher affinity for the protein than the corresponding peptoid, which is in

agreement with expectations.[10] It should be noted that this effect was most pronounced in case of the viral protease. The binding of peptides to proteases has been shown in many cases to be largely dependent on the binding of the peptide backbone.[15] This is one of the reasons for the modest specificity of many proteases. The preference for retropeptoids over directly translated peptoids may thus be highly dependent on the receptor system of interest. In essence, the hypothesis regarding the action of peptoids is still based on two-dimensional (2D) similarity.

Substance P and the Neurokinin-1 (NK_1) Receptor System

Substance P peptoid analogues have been shown to be active as agonist mimics of substance P (SP) at the murine NK_1 receptor.[22,23] Both the peptoid and retropeptoid of the entire undecapeptide (Fig. 1, **B** and **C**, respectively), as well as the hexa(retro)peptoids for the C-terminal minimal sequence of substance P, displayed full agonist activity, albeit with moderate potencies.[22,23] Here, only a brief overview of the NK_1 tachykinergic receptor system is presented. For extensive reviews on this system see refs. 24-27. SP was first isolated in 1931 by von Euler and Gaddum from horse brain and intestine.[28] In later years another four mammalian tachykinins were isolated.[27] The name *tachy*kinin was given to these compounds in relation to their rapid stimulatory action on extravascular smooth muscle; *tachus* is derived from Greek and means fast. Not until forty years after its discovery, however, in 1971 the amino-acid sequence of SP was determined (Table 1).[29] Comparison of the sequences of the five mammalian tachykinins known today[27] reveals a common C-terminal sequence, Phe-Xxx-Gly-Leu-Met-NH_2 in which Xxx is a hydrophobic residue. This consensus sequence is characteristic for tachykinins and is mainly responsible for their biological activity at the neurokinin receptors. Up to ten additional tachykinins have been isolated from lower species (mainly amphibians) as well.[30,31] The five *mammalian* tachykinins are specifically referred to as neurokinins.[27]

SP is released from sensory nerve endings throughout the body.[26,27] *In vivo* SP is rapidly degraded and deactivated by various proteolytic enzymes, such as neutral endopeptidase (NEP) or angiotensin converting enzyme (ACE).[27] The presence of the C-terminal amide group in neurokinins precludes degradation by carboxy endopeptidases,[26,27] and also artificial methylation of the amide nitrogens in SP analogues has shown to result in improved resistance towards a SP hydrolyzing enzyme.[32] In mammals three receptor

subtypes for the neurokinins exist, the NK_1 (neurokinin-1), NK_2 and NK_3 receptor. The neurokinins bind to and activate all three receptor subtypes with different selectivity; SP is the preferred ligand for the NK_1 receptor.

TABLE 1. Amino-acid sequences of the mammalian tachykinins (neurokinins)[26,27]

Name		Sequence
Substance P	SP	H-Arg-Pro-Lys-Pro-Gln-Gln-**Phe**-Phe-**Gly-Leu-Met-NH_2**
Neurokinin A	NKA	H-His-Lys-Thr-Asp-Ser-**Phe**-Val-**Gly-Leu-Met-NH_2**
Neurokinin B	NKB	H-Asp-Met-His-Asp-Phe-**Phe**-Val-**Gly-Leu-Met-NH_2**
Neuropeptide K[a]	NPK	H-Asp-Ala-Asp-Ser-Ser-Ile-Glu-Lys-Gln-Val-Ala-Leu-Leu-Lys-Ala-Leu-Tyr-Gly-His-Gly-Gln-Ile-Ser-His-Lys-Arg-His-Lys-Thr-Asp-Ser-**Phe**-Val-**Gly-Leu-Met-NH_2**
Neuropeptide γ[a]	NPγ	H-Asp-Ala-Gly-His-Gly-Gln-Ile-Ser-His-Lys-Arg-His-Lys-Thr-Asp-Ser-**Phe**-Val-**Gly-Leu-Met-NH_2**

The tachykinin consensus sequence is given in bold. a) Neuropeptide K (NPK) and neuropeptide γ (NPγ) are N-terminal extended forms of neurokinin A (NKA).

The three neurokinin receptor subtypes are G-protein coupled receptors that belong to the rhodopsin superfamily.[33] The receptors are integral membrane proteins consisting of seven more or less parallel helical domains that span the membrane connected by intra- and extracellular loops. Although an experimentally determined three-dimensional (3D) structure of a G-protein coupled receptor is still lacking, it is possible to construct crude models of the transmembrane part of the receptor structure.[33] A schematic representation of the human NK_1 receptor is depicted in Fig. 2. The ligand binding sites are mainly formed by amino-acid side chains of the transmembrane (TM) domains and extracellular loops that are in close proximity in the 3D model.[37] The agonist binding site probably consists of multiple epitopes located in the extracellular N-terminus and the extracellular loops, whereas the binding sites for non-peptide antagonists are mainly formed by the 'top' parts of TM-III, IV, V and VI (Fig. 2).[37] Previously, the activation mechanism of G-protein coupled receptors was solely believed to be a conformational change in the receptor structure induced by binding of the agonist. Recently, the picture emerges that the 'structure' of the receptor is in fact a number of interconverting receptor conformations while maintaining the overall seven TM helical structure. The agonist is thereby believed to stabilize the active conformation which triggers the G-protein and the subsequent signal-transduction cascade.[38,39]

The neurokinins and their receptors have been reported to play a role in neuronal development,[40] in the transmission of pain[25] and e.g. in airway hyperresponsiveness such as occurs in allergic asthma.[41] The inflammatory responses that often occur upon the release of neurokinins from sensory nerve endings are generally referred to as neurogenic inflammation. Antagonists for the NK_1 and NK_2 receptors may therefore play a therapeutic role as anti-inflammatory compounds. In addition, they may also play a role as analgesics in the treatment of chronic pain or as anti-emetic compounds, that may reduce nausea as a result of e.g. chemotherapy. The full physiological and pathophysiological role of neurokinins in these and many other processes[25,41] is, however, far from clear.

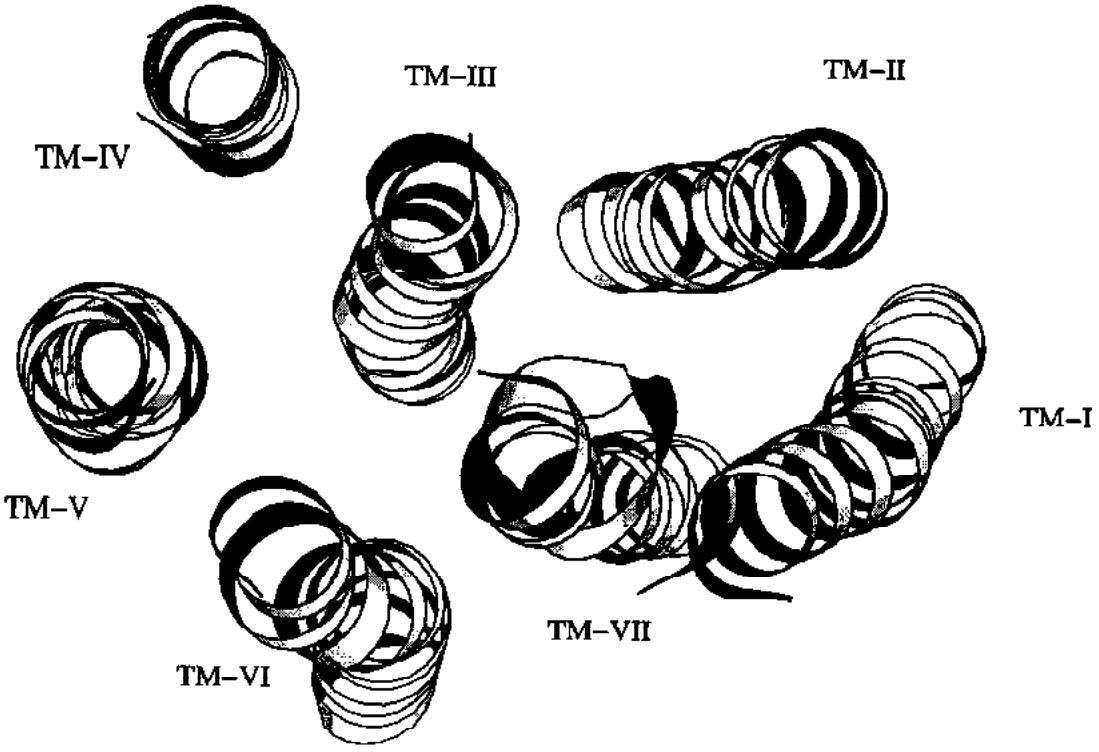


FIGURE 2. A model of the helical transmembrane (TM) domains of the human NK_1 receptor seen in projection from the extracellular ('top') side. The plane of the paper is parallel to the surface of the cell membrane. The helices are depicted using MOLSCRIPT[34] and are given in an arrangement derived from the projection structure of bovine rhodopsin.[35,36]

The NK_1 antagonists CP-99,994 (**2**) and FK888 (**7**) have been evaluated in clinical trials concerning asthma. FK888 improved the recovery from airway narrowing (bronchoconstriction) in asthmatic patients that was induced by physical exercise,[42] on the other hand CP-99,994 failed to reduce this effect after inhalation of multiple doses of hypertonic saline of increasing concentration.[43]

Bioactive Conformation of Substance P

Many modifications of SP have been synthesized in the past to study its structure-activity relationship (SAR),[44] and to find peptide antagonists that can be used as pharmacological tools. SP can in a stepwise fashion be shortened starting at the N-terminus without a severe loss of agonist activity; the minimal sequence of SP that is necessary for NK_1 agonist activity is the C-terminal hexapeptide $SP_{6\text{-}11}$ (Fig. 3).[44] The N-terminal part of the sequence, on the other hand, determines the selectivity for the NK_1 receptor.[27,45] The minimal sequence $SP_{6\text{-}11}$ contains one N-terminal amino-acid residue (Gln^6) in addition to the tachykinin consensus sequence (see Table 1).[#] The glutamine residue in position 6 (Gln^6) can be replaced by other hydrophilic residues or groups that contain either a basic amino group or an amide side chain, such as arginine (Arg), δ-amino valeric acid (Ava), which is a lower homologue of lysine, or pyroglutamic acid (pGlu).[44] The two phenylalanines (Phe^7 and Phe^8) directly following Gln^6 are crucial for activity; a change of chirality at Phe^7 and/or Phe^8 results in a nearly complete loss of agonist activity.[44] Efforts to determine the bioactive side-chain conformations of these two phenylalanines using rigidization gave conflictual results, which have been attributed to complex ligand-receptor dynamics.[46]

The succeeding glycine residue (Gly^9) in the peptide chain is a site of considerable flexibility in the SP backbone (Fig. 3). Restricting the conformation of Gly^9 to various parts of the Ramachandran plot,[16] in peptides derived from $SP_{6\text{-}11}$, determines the selectivity for the three receptor subtypes.[27] Selectivity for the NK_1 receptor over the other two receptors is achieved by substituting L-Pro for Gly^9 in $SP_{6\text{-}11}$ or by methylation of the amide nitrogen of Gly^9 to yield a sarcosine residue (Sar).[27] Replacing Gly^9 by D-Pro, on the other hand, favours selectivity for the NK_2 receptor in rats.[47] Furthermore, restricting both the ϕ and ψ backbone torsion angles of Gly^9 by means of an (*R*)- or (*S*)-spirolactam fragment (Fig. 3)

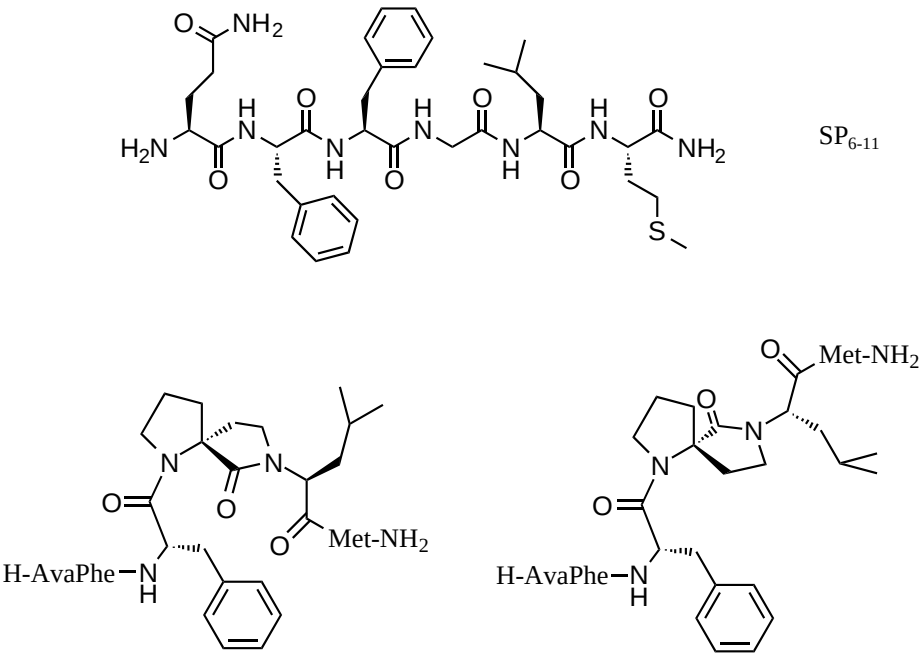
[#] Also in shortened analogues of SP the amino-acid residues are indicated with respect to their position in $SP_{1\text{-}11}$.

renders [Ava^6]SP$_{6-11}$ into an NK$_1$ agonist and antagonist, respectively.[48] In view of these findings, it has been concluded that NK$_1$ agonism or antagonism of SP$_{6-11}$ is related to conformations of Gly^9 in the ϕ,ψ (-,+) or ϕ,ψ (+,-) quadrant of the Ramachandran plot, respectively.[48] Meanwhile, NK$_2$ agonism is preferred when Gly^9 is replaced by a cyclic constraint which restricts the conformation of this residue to the ϕ,ψ (+,+) quadrant.[27] It has been proposed that in solution an (*S*)-spirolactam ring on position 9 in [Ava^6]SP$_{6-11}$ stabilizes a β-turn.[48] On the other hand, a linear rather than a turn conformation has been proposed for an (*R*)-spirolactam ring in SP$_{6-11}$ based on the ^{1}H NMR spectrum in DMSO-d_6 in combination with additional conformational energy calculations.[44] Preliminary studies[49] revealed, however, that it cannot be excluded that the (*R*)-spirolactam constraint may stabilize a β-turn as well.

Oxidation of the sulfur atom in the Met^{11} side chain to either a sulfoxide or a sulfone increases selectivity of SP$_{1-11}$ and SP$_{6-11}$ for the NK$_1$ receptor.[50] Conformationally constrained residues at position 9 and/or oxidation of Met^{11} have resulted in the selective NK$_1$ agonists septide ([pGlu^{6}Pro^9]SP$_{6-11}$) and [Sar^{9}Met(O$_2$)^{11}]SP.[50,51] It should, however, be noted that different binding modes for agonist peptides derived from either SP$_{6-11}$ or SP have been suggested.[52,53] In addition to these local conformational constraints, the entire minimal sequence of SP can be incorporated in ring structures while maintaining agonist activity.[54] These global conformational constraints restrict the number of energetically accessible conformations of the peptide, though considerable conformational flexibility remains. Despite the use of cyclic SP analogues[54] and the knowledge gained with respect to the bioactive conformation of Gly^9, the bioactive conformation of SP largely remains in the dark.

Non-Peptide NK$_1$ Antagonists

Historically, most NK$_1$ antagonists have been obtained by making modifications of SP, that thereby largely maintain a peptidic character. Perhaps as a result, these compounds are in most cases not very potent antagonists.[27] On the other hand, most of the highly potent NK$_1$ antagonists were obtained by the optimization of non-peptide compounds that had been identified by random screening.[e.g. 55,56] High-affinity (non)peptide NK$_1$ antagonists contain at least two aromatic rings, that are held in a proper position by a great number of possible scaffolds.[57] The antagonists may be classified based on these scaffolds.



[Ava^{6}-'(*R*)spirolactam'^{9}]SP$_{6-11}$ [Ava^{6}-'(*S*)spirolactam^{9}]SP$_{6-11}$

ϕ,ψ (-42,138) ϕ,ψ (42, -138)

FIGURE 3. The minimal sequence SP$_{6-11}$ of substance P and two analogues with conformational constraints at the position of glycine residue no. 9. The (*R*)-spirolactam analogue is an NK$_1$ agonist; the (*S*)-spirolactam analogue is an NK$_1$ antagonist.

Some antagonists, as representatives of an entire antagonist class are depicted in Fig. 4: quinuclidine based antagonists (**1**, CP-96,345),[56,58] that appear to be closely related to the piperidine based antagonists (**2**, CP-99,994),[59] perhydroisoindole antagonists (**3**, RP67580),[60,61] L-tryptophan benzylesters (**4**, L-732,138),[62] SR140333 (**5**),[63] isoquinoline-urea and pyrido[3,4-*b*]pyridine carboxamide antagonists (**6**),[64] dipeptide based antagonists (**7**, FK888),[65] diacylpiperazines (**8**, L-161,664),[66] and antagonists containing a steroid skeleton (**9**, WIN51708).[67] The selection of antagonists depicted in Fig. 4 is restricted to those antagonists for which the receptor binding site has been explored using site-directed mutagenesis, or for which a crystal structure is available. Only for compound **6** binding site data has not been reported.

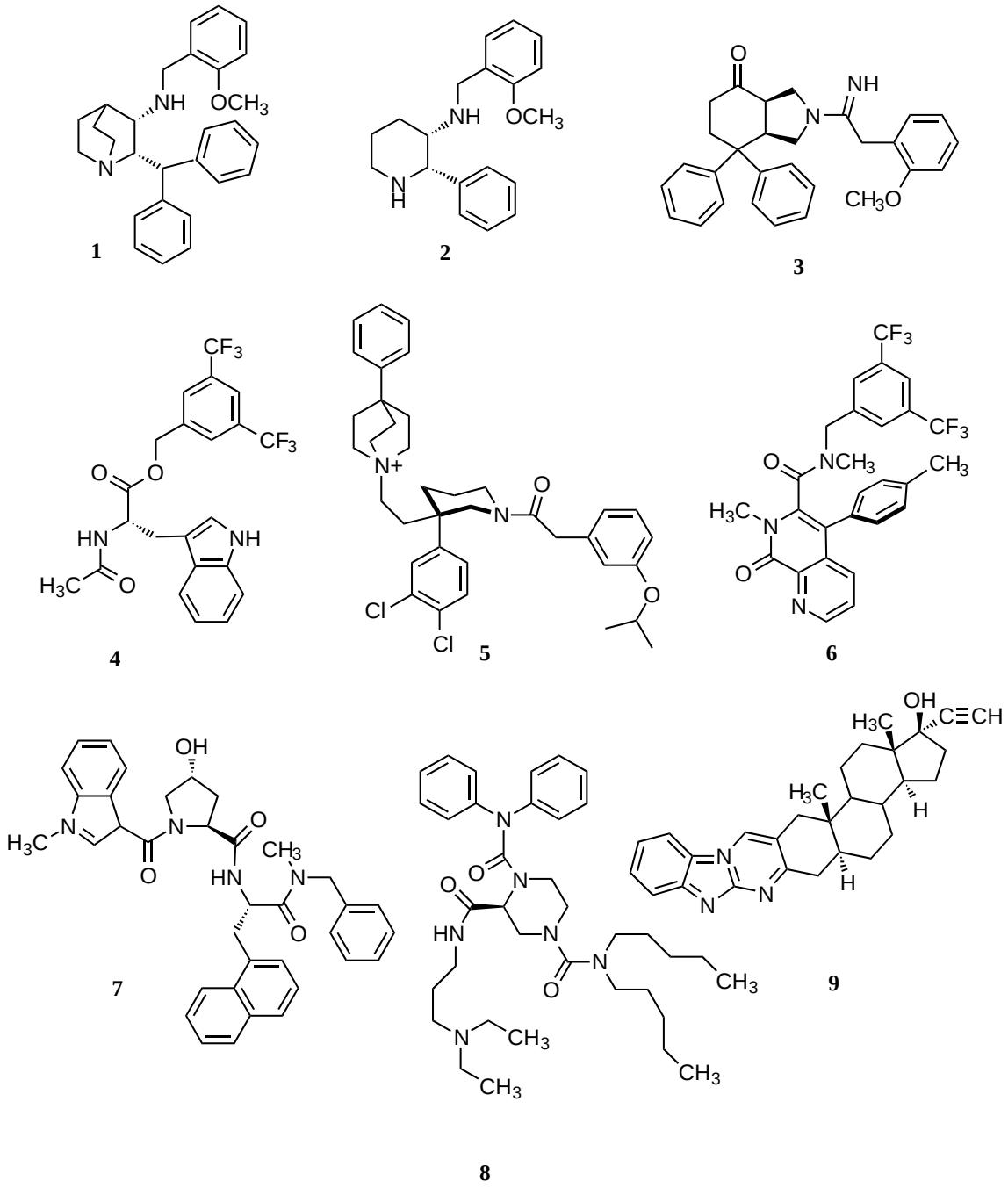
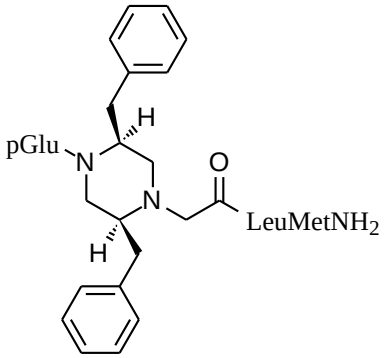
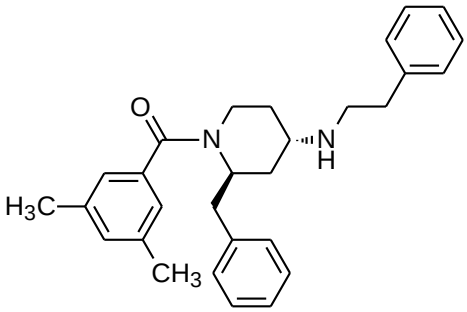


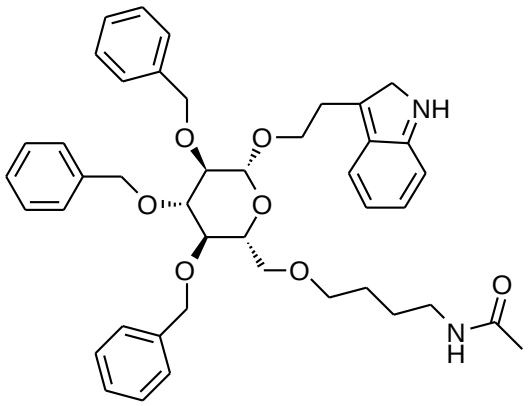
FIGURE 4. A selection of (non)peptide NK_1 antagonists as representatives of the various antagonist classes.



10



11



12

FIGURE 5. NK_1 antagonists resulting from a peptidomimetic approach applied to SP (resulting in **10**, **11**) and somatostatin (resulting in **12**).

A small number of (non)peptide antagonists has resulted from a more rational approach exploiting resemblance to SP. This has resulted in both a less peptidic but also very low potency piperazine based antagonist **10**,[68] and in the more potent antagonist CGP47899 (**11**) (Fig. 5).[69] Remarkably, a glucose based antagonist that initially had been designed to mimic the putative β-turn in the $Phe^7Trp^8Lys^9Thr^{10}$ sequence of the hormone somatostatin, turned out to be a relatively potent SP antagonist.[6]

The importance of the two adjacent phenylalanine residues in the SP sequence (Phe^7 and Phe^8) and the importance of aromatic rings of the high-affinity NK_1 antagonists has led to the 'phenyl-phenyl hypothesis': The aromatic rings in the NK_1 antagonists might block the binding sites for the phenyl side chains of Phe^7 and Phe^8, thereby preventing the binding of

SP.[69] For many of the non-peptide NK_1 antagonists, however, binding sites that are distinct from the SP binding site have been reported.[70,71] Meanwhile, their competitive antagonism (**5** is non-competitive)[63] is supposed to be carried into effect by an allosteric mechanism.[71] Despite the considerable diversity in antagonist scaffolds the diversity in binding sites at the human NK_1 receptor seems to be limited;[71] even for the structurally very distinct benzylether analogue L-709,210 of the benzylamino quinuclidine antagonist **1**, and the L-tryptophan benzylester **4** similar binding sites have been reported.[72] On the other hand, a pronounced selectivity of e.g. CP-96,345 (**1**) and RP67580 (**3**) for the NK_1 receptors in different species has been reported.[73] In the antagonist binding sites a considerable number of aromatic amino-acid residues has been identified.[71] This seems to be consistent with the importance of aromatic rings in the NK_1 antagonists (Fig. 4), the important role attributed to aromatic-aromatic interactions in the stabilization of protein structures[74] and in molecular recognition phenomena.[75] It has been suggested that the diphenylmethyl group (benzhydryl group) of the quinuclidine antagonists (**1**) is either surrounded by a cluster of aromatic amino-acid residues (His-197, Phe-268 and Tyr-272)[76] that is stabilized by Glu-193 and Lys-194, or is mainly bound by a histidine residue (His-197) via amino-aromatic interactions.[77] The proposed mutual proximity of the aromatic residues and Glu-193 and Lys-194 in the human NK_1 receptor has been supported by the possibility to create a Zn^{2+} binding site in the receptor by gradually mutating a number of these residues into histidine residues;[78] binding of SP to this multiple histidine mutant was virtually unchanged compared to the wild-type receptor.[78]

Molecular Recognition

Peptoids have been developed to mimic the action of peptides.[10,13,22] The explanation for this mimicry is, however, still based on the topological similarity of peptides and peptoids.[10] Understanding the mechanism of action of peptoids includes describing the structure and the conformational behaviour of peptoids in general and developing a theory regarding the mechanism of molecular recognition. This would include postulating models for the bioactive conformation of the substance P peptoids and giving a description of the intermolecular interactions involved. Irrespective of the activation mechanisms of agonists or the inhibiting mechanisms of antagonists at the receptor level, the first step always involves binding of a ligand (L) to its receptor (R):

$$L + R \rightleftharpoons [LR]_{complex}$$

The selectivity of this binding process and the molecular mechanisms behind it make up what is called molecular recognition. The definition of peptidomimetics as being compounds that can mimic the critical features of the molecular recognition process of the parent peptide, has to be translated into workable concepts to direct research in this field. The following three concepts may serve this purpose. (i) *Complementarity*. Between the ligand and the binding site in the complex there will be complementarity of shape and electrostatic properties to a certain extent. (ii) *Similarity*. Different ligands that bind at the same site of the receptor probably have critical features of their interactions in common; this similarity between an exogenous compound and the natural ligand is here referred to as mimicry. (iii) *Selectivity*. Only specific compounds can function as a ligand for a particular receptor, and not every modification of a ligand binds. Alternatively, the binding process can be approached from a more fundamental point of view, i.e. in terms of enthalpic and entropic changes for both ligand and receptor upon complexation.[79] The three concepts mentioned can, however, more directly be translated into protocols for the computer-aided study of molecular recognition:

In studying *complementarity* the ligand-receptor complex itself is the central issue. In the ideal case experimentally determined ligand-receptor complexes are available for the system of interest. For the tachykinergic receptor system this information unfortunately is lacking. Yet, crystallographically determined peptide-protein[15,4] and peptidomimetic-protein[e.g.80] complexes can be used for understanding the nature of the electrostatic[81] and steric complementarity[82,83] involved in binding. Furthermore, experimentally determined protein-ligand complexes may form the basis for the construction and validation of theoretical methods to study the selectivity and similarity of ligands. An extensive and thorough review on the use of protein-ligand complexes in understanding molecular recognition and for improving ligand design appeared recently.[84] When an experimentally determined structure of the receptor is lacking, but sufficient sequence homology between the target receptor and known macromolecular structures exists, a homology model of the receptor can be constructed.[85] Docking of (possible) ligands to receptor structures or receptor models can, in addition, be a powerful tool that is also based on ligand-receptor complementarity.[86,87]

Selectivity and *similarity* of compounds can at a theoretical level be studied by tools that are often employed in indirect ligand design.[88] In cases where only the structures of the ligands are known (i.e. indirect ligand design) one can turn to studying the similarity of properties that describe or determine intermolecular interactions.[89,90] Such work often

includes the construction of a three-dimensional arrangement of functional groups that is common to molecules that share the same binding site, i.e. a pharmacophore.[e.g. 91] In designing peptide turn-mimetics, pharmacophore-like descriptions of the conformation of a peptide backbone[92] or of side-chain directions[93] have for instance been employed. In more general terms descriptions of the similarity of intermolecular interactions may include e.g. common features of the molecular electrostatic potential,[94] the shape of the ligands[95] or the solvent accessibility of functional groups. *Selectivity* of a pharmacophore can e.g. be validated by searching databases containing three-dimensional molecular structures and their known pharmacological activities.[96] It is important to note that the notion of a pharmacophore need thereby not be restricted to the nature of functional groups, but can be extended to so-called 'interactional properties', such as the molecular electrostatic potential.[97] Dockings of possible ligands to a known active site and subsequently estimating the Gibbs free-energy of binding is on the other hand also a means of studying selectivity.[98,99]

Conformational Behaviour of Ligands

For studying the molecular recognition of peptoids, valid three-dimensional models of peptoids are required. Understanding the conformational behaviour of peptoids, e.g. knowledge of the preferred as well as the energetically inaccessible conformations and the influence of the molecular environment is thereby a basic step; molecular properties are dependent on the structure (and thus the conformation) of a molecule. The conformational behaviour of *peptides* has been found to be very dependent on the environment.[100] For instance, SP is very flexible in aqueous solution, but clearly preferred conformations are induced when the molecular environment becomes more hydrophobic.[27] This stresses the importance of understanding the conformational behaviour of peptides as well as peptoids *at a protein binding site*, e.g. by identifying the conformations of the backbone and of the side chains in the constituting monomers when bound to a protein. Ultimately, we want to determine the bioactive conformation of the entire molecule. The conformational behaviour of peptides and peptoids in solution can be studied by means of NMR techniques.[101] Energetically accessible conformations are given by crystal structures of the non-complexed ligands.[102] Neither of these two are, however, sufficient to locate the bioactive conformation or to study the ligand's behaviour at protein binding sites. To study aspects of molecular recognition in a system of interest that is not directly under experimental scrutiny, the use of theoretical methods is then the only possibility. In that respect, it is important to note that

bioactive conformations (as determined in protein-ligand complexes) in a considerable number of cases have been found to differ from the global minimum energy conformation by conformational energies well up to ca. 30 kcal/mol on the *in vacuo* energy scale.[103] This 'energy penalty' depends on the requirements of the protein binding site, but also on the number of freely rotatable bonds and the size of the ligand.[103] A large number of rotatable bonds such as is the case of peptoids therefore forces one to consider a large number of energetically accessible conformations when in search for the unknown bioactive conformation. Furthermore, the mode of binding as well as the bioactive conformation may change dramatically as a result of only small modifications in the ligand.[104,84]

With respect to the study of the molecular mechanism of action of non-macromolecular bioactive compounds, crystal structures of these so-called 'small molecules' play an important role. Small-molecule crystal structures yield energetically accessible conformations and intermolecular interactions in a highly structured environment.[102,105] The intermolecular interaction geometries in small-molecule crystal structures and in ligand-receptor complexes have been found to reveal the same geometrical preferences.[106] In addition, the conformational preferences of fragments have been found to be similar in small-molecule crystal structures and in ligands in protein binding sites.[107] In conclusion, small-molecule crystal structures as well as crystal structures of protein-ligand complexes continue to be valuable reference points for empirical theoretical methods that are aimed at studying and describing molecular recognition.

Scope of this Thesis

The work described in this thesis mainly addresses the structure and conformational behaviour of peptoid peptidomimetics. The tachykinergic NK_1 receptor system and its agonists and antagonists thereby mainly serve as vehicles; a thorough understanding of the molecular mechanism of action of non-peptide NK_1 antagonists and especially of the substance P peptoids is the distant goal. Research in this direction is mainly restricted to the study of the ligands, since an experimentally determined structure of the (human) NK_1 receptor is yet unknown. Emphasis in this work is therefore directed to the largely unexplored field of the structure and conformational behaviour of peptoids and to the study of methods to investigate intermolecular interaction geometries. Insight in structures and interactions is a necessary basis for studying molecular recognition and mimicry. This knowledge is also of a more general importance in understanding the mechanism of action of

peptoids in general, and is of relevance for interpreting the results of pharmacological tests on series of (SP) peptoid analogues. Those studies should facilitate the proposal of new compounds with preferably a higher affinity, efficacy and hopefully a predefined selectivity.

Based on crystal structure data for a number of non-peptide NK_1 antagonists, the intermolecular interactions of these antagonists with their environment were evaluated and interaction geometries with amino-acid residues in the human NK_1 receptor are proposed (Chapter 2). To be able to expand the study of interactions of compounds to any selected conformation and validly predict interaction geometries, each other complementing, recently developed empirical methods are evaluated (Chapter 3). Such methods are highly required to be able to study and describe mimicry of the molecular recognition process of peptides and peptidomimetics, independent of the molecular framework.

Given the need for sound structural knowledge on the parent peptide and its mimicking peptoids, attempts were made to crystallize substance P and its retropeptoid and determine their crystal structures (Chapter 4). So far, these attempts did not yield the desired results. In addition, the geometrical and conformational preferences of characteristic peptoid fragments in small-molecule crystal structures and in protein binding sites are studied (Chapters 5 and 6). The knowledge gained here is used to assess the reliability and applicability of theoretical conformational analysis on peptoid oligomers (Chapter 7) and to make the first step to understand the molecular basis of mimicry by peptoids. Implications of the work described in this thesis are given in Chapter 8.

REFERENCES

1. Hökfelt, T., Johansson, O., Ljungdahl, Å., Lundberg, J. M., Schultzberg, M., *Nature*, **1980**, *284*, 515.
2. Witkop, B., *Med. Res. Rev.*, **1992**, *12*, 275.
3. Roitt, I. M., Brostoff, J. and Male, D. K., *Immunology*, 3rd. ed., Mosby, London, 1993.
4. Wilson, I. A. and Stanfield, R. L., *Curr. Opin. Struct. Biol.*, **1993**, *3*, 113.
5. Van den Elsen, J. M. H., *Antibody Recognition of Neisseria meningitidis*, thesis, Utrecht University, 1996.
6. Giannis, A. and Kolter, T., *Angew. Chem. Int. Ed. Engl*., **1993**, *32*, 1244.
7. Taylor, M. D. and Amidon, G. L., *Peptide-Based Drug Design. Controlling Transport and Metabolism*, American Chemical Society, Washington, 1995.
8. Balkenkohl F., Von dem Bussche-Hünnefeld, C., Lansky, A. and Zechel, C., *Angew. Chem. Int. Ed. Engl.*, **1996**, 35, 2288.
9. Liskamp, R. M. J., *Angew. Chem. Int. Ed. Engl.*, **1994**, *33*, 633.
10. Simon, R. J., Kania, R. S., Zuckermann, R. N., Huebner, V. D., Jewell, D. A., Banville, S., Ng, S., Wang, L., Rosenberg, S., Marlowe, C. K., Spellmeyer, D. C., Tan, R.,

Frenkel, A. D., Santi, D. V., Cohen, F. E. and Bartlett, P. A., *Proc. Natl. Acad. Sci. USA*, **1992**, *89*, 9367.
11. Kessler, H., *Angew. Chem. Int. Ed. Engl.*, **1993**, *32*, 543.
12. IUPAC-IUB Commission on Biochemical Nomenclature, *J. Mol. Biol.*, **1970**, *52*, 1.
13. Zuckermann, R. N., Martin, E. J., Spellmeyer, D. C., Stauber, G. B., Shoemaker, K. R., Kerr, J. M., Figliozzi, G. M., Goff, D. A., Siani, M. A., Simon, R. J., Banville, S. C., Brown, E. G., Wang, L., Richter, L. S. and Moos, W. H., *J. Med. Chem.*, **1994**, *37*, 2678.
14. Martin, E. J., Blaney, J. M., Siani, M. A., Spellmeyer, D. C., Wong., A. K. and Moos, W. H., *J. Med. Chem.*, **1995**, *38*, 1431.
15. Stanfield, R. L., Wilson, I. A., *Curr. Opin. Struc. Biol.*, **1995**, *5*, 103.
16. Ramachandran, G. N. and Sasisekharan, V., *Adv. Protein Chem.*, **1968**, *23*, 283.
17. Karle, I. L., *Acta Cryst.*, **1992**, *B48*, 341.
18. Kruijtzer, J. A. W. and Liskamp, R. M. J., *Tetrahedron Lett.*, **1995**, *36*, 6969.
19. Kessler, H., Bermel, W., Müller, A. and Pook, K.-H., *The Peptides*, Vol. 7, Chapter 9, Academic Press, 1985.
20. Moehle, K. and Hofmann, H.-J., *Biopolymers*, **1996**, *38*, 781.
21. Moehle, K. and Hofmann, H.-J., *J. Mol. Model.*, **1996**, *2*, 307.
22. Westra-de Vlieger, J. F., Van Heuven-Nolsen, D, Wilting, J., Kruijtzer, J. A. W., Liskamp, R. M. J. and Nijkamp, F. P., *in preparation*, **1997**.
23. Kruijtzer, J. A. W., *Synthesis of Peptoid Peptidomimetics*, thesis, Utrecht University, 1996.
24. Pernow, B., *Pharmacol Rev.*, **1983**, *35*, 85.
25. Otsuka, M. and Yoshioka, K., *Physiol. Rev.*, **1993**, *73*, 229.
26. Maggi, C. A., Patachini, R., Rovero, P. and Giachetti, A., *J. Auton. Pharmacol.*, **1993**, *12*, 23.
27. Regoli, D., Boudon, A. and Fauchere, J.-L., *Pharmacol. Rev.*, **1994**, *46*, 551.
28. Euler, U. S. von, and Gaddum, J. H., *J. Physiol. (Lond.)*, **1931**, *72*, 74.
29. Chang, M. M., Leeman, S. E. and Niall, H. D., *Nature*, **1971**, *232*, 86.
30. Ersparmer, V. and Melchiorri, P., *Trends Pharmacol. Sci.*, **1980**, *1*, 391.
31. Mussap, C. J., Geraghty, D. P. and Burcher, E., *J. Neurochem.*, **1993**, *60*, 1987.
32. Sandberg, B. E. B., Lee, C.-M., Hanley, M. R., and Iversen, L. L., *Eur. J. Biochem.*, **1981**, *114*, 329.
33. Donnelly, D. and Findlay, J. B. C., *Curr. Opin. Struc. Biol.*, **1994,** *4*, 582; Baldwin, J. M., *Curr. Opin. Struc. Biol.*, **1994**, *6*, 180.
34. Kraulis, P. J., *J. Appl. Cryst.*, **1991**, *24*, 946.
35. Schertler, G. F. X., Villa, C. and Henderson, R., *Nature*, **1993**, *362*, 770.
36. Bhogal, N., Donnelly, D. and Findlay, J. B. C., *J. Biol. Chem.*, **1994**, *269*, 27269; Donnelly, D., Findlay, J. B. C. and Blundell, T. L., *Receptors and Channels*, **1994**, 2, 61.
37. Schwartz, T. W., *Curr. Opin. Biotech.*, **1994**, *5*, 434.
38. Kenakin, T., *Trends Pharmacol. Sci.*, **1995**, *16*, 188.
39. Rosenkilde, M. M., Cahir, M., Gether, U., Hjorth, S. A. and Schwartz, T. W., *J. Biol. Chem.*, **1994**, *269*, 28160.
40. De Felipe, C., Pinnock, R. D. and Hunt, S. P., *Science*, **1995**, *267*, 899.
41. Bertrand, C. and Geppetti, P., *Trends Pharmacol. Sci.*, **1996**, *17*, 255.

42. Ichinose M., Miura, M., Yamauchi, H., Kageyama, N., Tomaki, M., Oyake, T., Ohuchi, Y., Hida, W., Miki, H., Tamura, G. and Shirato, K., *Am .J. Respir. Crit. Care Med.*, **1996**, *153*, 936.
43. Fahy, V. J, Wong, H. H., Geppetti, P., Reis, J. M., Harris, S. C., Maclean, D. B., Nadel, J. A. and Boushey, H. A., *Am .J. Respir. Crit. Care Med.*, **1995**, *152*, 879.
44. Dutta, A. S., *Drugs of the Future*, **1987**, 781.
45. Schwyzer, R., *Biopolymers*, **1995**, *37*, 5; Schwyzer, R., *EMBO J.*, **1987**, *6*, 2255.
46. Josien, H., Lavielle, S., Brunissen, A., Saffroy, M., Torrens, Y., Beaujouan, J.-C., Glowinski, J. and Chassaing, G., *J. Med. Chem.*, **1994**, *37*, 1586; Josien, H., Convert, O., Berlose, J.-P., Sagan, S., Brunissen, A., Lavielle, S. and Chassaing, G., *Biopolymers*, **1996**, *39*, 133; Déry, O., Josien, H., Grassi, J., Chassaing, G., Couraud, J. Y. and Lavielle, S., *Biopolymers*, **1996**, *39*, 67.
47. Deal, M. J., Hagan, R. M., Ireland, S. J., Jordan, C. C., McElroy, A. B., Porter, B., Ross, B. C., Stephens-Smith, M. and Ward, P., *J. Med. Chem.*, **1992**, *35*, 4195.
48. Ward, P., Ewan, G. B., Jordan, C. C., Ireland, S. J., Hagan, R. M. and Brown, J. R., *J. Med. Chem.*, **1990**, *33*, 1848.
49. Boks, G. J., *unpublished results*.
50. Regoli, D., Drapeau, G., Dion S. and Couture, R., *Trends Pharmacol. Sci.*, **1988**, *9*,
290; Tousignant, C., Guillemette, G., Drapeau, G. and Regoli, D., *Neuropeptides*, **1989**, *14*, 275.
51. Wormser, U., Laufer, R., Hart, Y., Chorev, M., Gilon, C. and Selinger, Z., *EMBO J.*, **1986**, *5*, 2805.
52. Glowinski, J., *Trends Pharmacol. Sci.*, **1995**, *16*, 365.
53. Huang, R. C., Huang, D., Strader, C. D. and Fong, T. M., *Biochemistry*, **1995**, *34*, 16467.
54. Grdadolnik, S. G., Mierke, D. F., Byk, G., Zeltser, I., Gilon, C. and Kessler, H., *J. Med. Chem.*, **1994**, *37*, 2145; Saulitis, J., Mierke, D. F., Byk, G., Gilon, C. and Kessler, H., *J. Am. Chem. Soc.*, **1992**, *114*, 4818.
55. Lowe, J. A., III, Drozda, S. E., Snider, R. M., Longo, K. P., Zorn, S. H., Morrone, J., Jackson, E. R., McLean, S., Bryce, D. K., Bordner, J., Nagahisa, A., Kanai, Y., Suga, O., Tsuchiya, M., *J. Med. Chem.*, **1992**, *35*, 2591.
56. Snider, R. M., Constatine, J. W., Lowe, J. A., III, Longo, K. P., Lebel, W. S., Woody, H. A., Drozda, S. E., Desai, M. C., Vinick, F. J., Spencer, R. W., Hess, H. J., *Science*, **1991**, *251*, 435.
57. Desai, M. C., *Exp. Opin. Ther. Patents*, **1994**, *4*, 315.
58. Swain, C. J., Seward, E. M., Cascieri, M. A., Fong, T. M., Herbert, R., MacIntyre, D. E., Merchant, K. J., Owen, S. N., Owens, A. P., Sabin, V., Teall, M., VanNiel, M. B., Williams, B. J., Sadowski, S., Strader, C. D., Ball, R. G., Baker, R., *J. Med. Chem.*, **1995**, *38*, 4793.
59. Desai, M. C., Lefkowitz, S. L., Thadeio, P. F., Longo, K. P., Snider, R. M., *J. Med. Chem.*, **1992**, *35*, 4911.
60. Garret, C., Carruette, A., Fardin, V., Moussaoui, S., Peyronel, J. F., Blanchard, J. C. and Laduron, P. M., *Proc. Natl. Acad. Sci. USA*, **1991**, *88*, 10208.

61. Fardin., V., Carruette, A., Menager, J., Bock, M., Flamand, O., Foucault, F., Heuillet, E., Moussaoui, S. M., Tabart, M., Peyronel, J. F. and Garret, C., *Neuropeptides*, **1994**, *26* (Suppl.), 34.
62. Macleod, A. M., Merchant, K. J., Cascieri, M. A., Sadowski, S., Ber, E., Swain, C. J., Baker, R., *J. Med. Chem.*, **1993**, *36*, 2044.
63. Emonds-Alt, X., Doutremepuich, J.-D., Heaulme, M., Neliat, G., Santucci, V., Steinberg, R., Vilain, P., Bichon, D., Ducoux, J.-P., Proietto, V., Van Broeck, D., Soubrié, P., Le Fur, G. and Brelière, J.-C., *Eur. J. Pharmacol.*, **1993**, *250*, 403.
64. Natsugari, H., Ikeura, Y., Kiyota, Y., Ishichi, Y., Ishimaru, T., Saga, O., Shirafuji, H., Tanaka, T., Kamo, I., Doi, T., Otsuka, M., *J. Med. Chem.*, **1995**, *38*, 3106.
65. Fujii, T., Murai, M., Morimoto, H., Maeda, Y., Yamaoka, M., Hagiwara, D., Miyake, H., Ikari, N. and Matsuo, M., *Br. J. Pharmacol.*, **1992**, *107*, 785.
66. Cascieri, M. A., Shaio, L., Mills, S. G., MacCoss, M., Swain, C. J., Yu, H., Ber, E., Sadowski, S., Wu.M. T., Strader, C. D and Fong, T. M., *Mol. Pharmacol.*, **1995**, *47*, 660.
67. Sachais, B. S. and Krause, J. E., *Mol. Pharmacol.*, **1994**, *46*, 122.
68. Chorev, M., Roubini, E., Gilon, C. and Selinger, Z, *Biopolymers*, **1991**, *31*, 725.
69. Schilling, W., Bittiger, H., Brugger, F., Criscione, L., Hauser, K., Ofner, S., Olpe, H.-R., Vassout, A. and Veenstra, S., In: *Perspectives in Medicinal Chemistry*, B. Testa, E. Kyburz, W. Fuhrer, R. Giger (eds.), VCA, Basel, 1993, p. 207-220.
70. Gether, U., Johansen, T. E., Snider, R. M., Lowe, J. R., III, Nakanishi, S. and Schwartz, T. W., *Nature*, **1993**, *362*, 345.
71. Schwartz, T. W., Gether, U., Schambye, H. T. and Hjorth, S. A., *Curr. Pharmaceut. Design.*, **1995**, *1*, 325.
72. Cascieri, M. A., Macleod, A. M., Underwood, D., Shiao, L. L., Ber, E., Sadowski, S., Yu, H., Merchant, K. J., Swain, C. J., Strader, C. D., Fong, T. M., *J. Biol. Chem.*, **1994**, *269*, 6587.
73. Pradier, L., Habert-Ortoli, E., Emile, L., Le Guern, J., Loquet, I., Bock., M.-D., Clot, J., Mercken, L., Fardin., V., Garret., C and Mayaux, J.-F., *Mol. Pharmacol.*, **1995**, *47*, 314; Jensen, C. J., Gerard, N. P., Schwartz, T. W. and Gether, U., *Mol. Pharmacol.*, **1994**, *45*, 294.
74. Burley, S. K. and Petsko, G. A., *Science*, **1985**, *229*, 23; Hunter, C. A., Singh, J. and Thornton, J. M., *J. Mol. Biol.*, **1991**, *218*, 837.
75. Hunter, C. A., *Chem Soc. Rev.*, **1994**, 101.
76. Gether, U., Lowe, J. A., III and Schwartz, T. W., *Biochem. Soc. Trans.*, **1995**, *23*, 96.
77. Fong, T. M., Cascieri, M. A., Yu, H., Bansal, A., Swain, C. and Strader, C. D., *Nature*, **1993**, *362*, 350.
78. Elling, C. E., Nielsen, S. M. and Schwartz, T. W., *Nature*, **1995**, *374*, 74.
79. Eftink, M. and Biltonen, R., In: *Biological Microcalorimetry*, A. E. Beezer (ed.), Academic Press, London, 1980, p. 343-412.
80. Lam, P. Y. S., Jadhav, P. K., Eyermann, C. J., Hodge, C. N., Ru, Y., Bacheler, L. T., Meek, J. L., Otto, M. J., Rayner, M. M., Wong, Y. N., Chang, C.-H., Weber, P. C., Jackson, D. A., Sharpe, T. R., and Erickson-Viitanen, S., *Science*, **1994**, *263*, 380.
81. Chau, P.-L. and Dean, P. M., *J. Comput.-Aided Mol. Design*, **1994**, *8*, 513.
82. Lawrence, M. C. and Colman, P. M., *J. Mol. Biol.*, **1993**, *234*, 946.

83. Laskowski, R. A., Thornton, J. M., Humblet, C. and Singh, J., *J. Mol. Biol.*, **1996**, *259*, 175.
84. Böhm, H.-J. and Klebe, G., *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 2588.
85. Blundell, T. L., Sibanda, B. L., Sternberg, M. J. E. and Thornton, J. M., *Nature*, **1987**, *326*, 347.
86. Kuntz, I. D., *Science*, **1992**, *257*, 1078.
87. Goodsell, D. S. and Olson, A. J., *Prot. Struc. Funct. Genet.*, **1990**, *8*, 195; Goodsell, D. S., Lauble, H., Stout, C. D. and Olson, A. J., *Prot. Struc. Funct. Genet.*, **1993**, *17*, 1; Stoddard, B. L. and Koshland, D. E., *Nature*, **1992**, *358*, 774.
88. Loew, G. H., Villar, H. O and Alkorta, I., *Pharmaceut. Res.*, **1993**, *10*, 475.
89. Rohrer, D. C., In: *Molecular Similarity and Reactivity: From Quantum Chemical to Phenomenological Approaches*, R. Carbo (ed.), Kluwer Academic Publishers, Dordrecht, 1995, p. 141-161.
90. Finn, P. W., In : *Molecular Modelling and Drug Design*, J. G. Vinter and M. Gardner (eds.), Macmillan Press Ltd., Hampshire, 1994, p. 266-304.
91. Desai, M. C., Lefkowitz, S. L., Thadeio, P. F., Longo, K. P., Snider, R. M., *J. Med. Chem.*, **1992**, *35*, 4911.
92. Olson, G. R., Bolin, D. R., Bonner, M. P., Bös, M., Cook, C. M., Fry, D. C., Graves, B. J., Hatada, M., Hill, D. E., Kahn, M., Madison, V. S., Rusiecki, V. K., Sarabu, R., Sepinwall, J., Vincent, G. P. and Voss, M. E., *J. Med. Chem.*, **1993**, *36*, 3039.
93. Horwell, D. C., Howson, W., Naylor, D. and Willems, H. M. G., *Bioorg. Med. Chem. Lett.*, **1995**, *5*, 1445.
94. Apaya, R. P., Lucchese, B., Price, S. L., and Vinter J. G., *J. Comput.-Aided Mol. Design*, **1995**, 9, 33; Van der Wenden, E. M., Price, S. L., Apaya, R. P., IJzerman, A. P. and Soudijn, W., *J. Comput.-Aided Mol. Design*, **1995**, 9, 44.
95. Arteca, G. A., In: *Reviews in Computational Chemistry*, Vol. 9, K. B. Lipkowitz, and D. B. Boyd (eds.), VCH Publishers, New York, 1996, p. 191-253.
96. Martin, Y. C., *J. Med. Chem.*, **1992**, *35*, 2145; Milne, G. W. A., Nicklaus, M. C., Driscoll, J. S. and Wang, S., *J. Chem. Inf. Comput. Sci.*, **1994**, *34*, 1219.
97. Thorner, D. A., Willett, P., Wright, P. M. and Taylor, R., *J. Comput.-Aided Mol. Design*, **1997**, *11*, 163.
98. Ajay and Murcko, M. A., *J. Med. Chem.*, **1995**, *38*, 4953.
99. Böhm, H.-J., *J. Comput.-Aided Mol. Design*, **1994**, *8*, 243.
100. Montcalm, T., Cui, W., Zhao, H., Guarnieri, F. and Wilson, S. R., *J. Mol. Struc.*, **1994**, *308*, 37.
101. Sadler, I. H., *Natural Prod. Rep.*, **1988**, 101.
102. Bürgi, H.-B., *Acta Cryst.*, **1988**, *B44*, 445.
103. Nicklaus, M. C., Wang, S., Driscoll, J. S. and Milne, G. W. A., *Bioorg. Med. Chem.*, **1995**, *3*, 411.
104. Meyer, E. F., Botos, I., Scapozza, L. and Zhang, D., *Persp. Drug Discov. Design*, **1995**, *3*, 168.
105. Brock, C. P. and Dunitz, J. D., *Chem. Mater.*, **1994**, *6*, 1118.
106. Klebe, G., *J. Mol. Biol.*, **1994**, *237*, 212.
107. Klebe, G. and Mietzner, T., *J. Comput.-Aided Mol. Design*, **1994**, *8*, 583; Klebe, G., *Persp. Drug Discov. Design*, **1995**, *3*, 85.

Chapter 2

Possible Ligand-Receptor Interactions for NK_1 Antagonists as Observed in Their Crystal Structures

Gertjan J. Boks[a], Jan P. Tollenaere[a] and Jan Kroon[b]

[a]*Department of Medicinal Chemistry, Utrecht Institute for Pharmaceutical Sciences, Universiteit Utrecht, Sorbonnelaan 16, NL-3584 CA Utrecht, The Netherlands*

[b]*Department of Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Universiteit Utrecht, Padualaan 8, NL-3584 CH Utrecht, The Netherlands*

ABSTRACT

Crystal structures of nine non-peptide tachykinin NK_1 antagonists have been analyzed for the intermolecular interactions of their pharmacophoric groups with neighbouring molecules in the crystals. Experimental data on interaction geometries of these antagonists with their environment can be of help in understanding the mechanism of binding to the human NK_1 receptor. Several N^+-aromatic and aromatic-aromatic interaction geometries for the positively charged quinuclidine rings and the pharmacophoric aromatic rings have been identified, which can explain the importance of aromatic amino-acid residues in the human NK_1 receptor. In addition, an interaction site for the side chain of Gln-165 in the human NK_1 receptor, that is probably involved in a hydrogen bond with the benzylamino nitrogen or benzylether oxygen of the quinuclidine and piperidine antagonists, is explicitly proposed. Also, a superposition based on pharmacophoric elements in the crystal structure conformations of two prototypic NK_1 antagonists, CP-96,345 and CP-99,994, suggests how both compounds might interact with the human NK_1 receptor in a similar manner.

Based on: Boks, G. J., Tollenaere, J. P. and Kroon, J., *Bioorg. Med. Chem.*, **1997**, *5*, 535-547

INTRODUCTION

Crystal structures have an established role in providing us with energetically accessible conformations of bioactive compounds and thereby provide a basis for structure-based ligand design.[1] For locating the possible intermolecular interactions of these compounds, however, one usually relies on information from secondary sources such as crystal structure statistics and theoretical studies on model systems. The information on interactions enclosed in the crystal structures of the compounds itself is scarcely used, although it may well reflect interactions with a macromolecular receptor.[1,2] Both in ligand-receptor binding and in crystallization a compromise is achieved between lowering the conformational energy and maximizing the number of intermolecular interactions. The underlying physical principles to both processes are the same and it is therefore likely that similar interaction geometries can be present both in the solid crystalline state of a ligand and when bound to a receptor. A detailed study of the crystal structures of nine tachykinin NK_1 antagonists was performed to document the experimentally observed interactions with their crystal environment. These interactions were then compared to knowledge on the bioactive conformations and the structure-activity relationships of the compounds, as well as to results from receptor mutagenesis studies that have been reported for some of these ligands.[3]

Non-peptide NK_1 antagonists are of interest, because of their potential anti-inflammatory and antinociceptive action. A possible role has been proposed for NK_1 antagonists in the relief of chronic pain and in the treatment of tachykinin mediated bronchoconstriction and plasma extravasation in asthma.[4] A dipeptide based NK_1 antagonist, FK888, has already entered clinical trials for the treatment of asthma.[5] In addition, a potent antiemetic NK_1 antagonist with good bioavailability was recently reported.[6] It might be valuable in the control of emesis (vomiting) resulting from, for example, chemotherapy. Detailed knowledge of the interaction mechanism of antagonists with the NK_1 receptor is of importance for the development of selective pharmacological tools. These can be helpful in further unravelling the role of tachykinins in a number of these pathophysiological processes and may open perspectives for the development of new drugs. In these efforts, our results can be used in the construction of binding-site or receptor models for these antagonists, in lack of an experimental structure of the human NK_1 receptor.

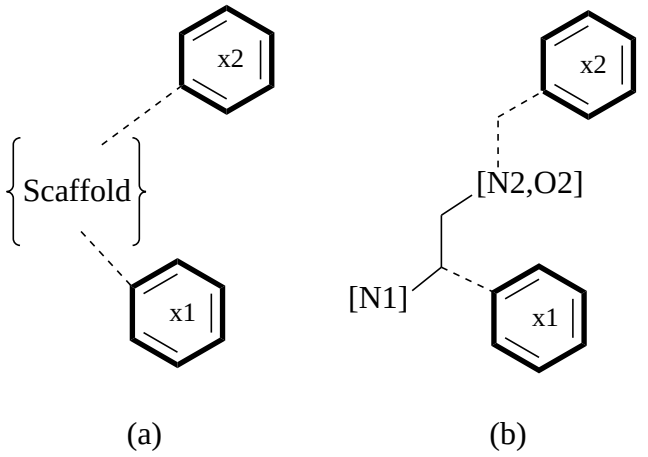


FIGURE 1. (a) Generalized non-peptide NK_1 antagonist pharmacophore consisting of two (or more) aromatic rings held together by various scaffolds. (b) A more detailed NK_1 non-peptide antagonist pharmacophore based on previously reported pharmacophoric elements for quinuclidine and piperidine NK_1 antagonists.

TABLE 1. Binding affinities for the human NK_1 receptor determined in various assays

No.	Name	IC_{50} (nM)	Ref.	No.	Name	IC_{50} (nM)	Ref.
1	CP-96,345	0.66 ± 0.26	14	**6**	CP-211,754	0.61 ± 0.058	10
2		12.2 ± 2.9	27	**7**	L-708,568	67 ± 10	26
3	L-709,210	0.2	12	**8**	L-732,244	22 ± 7	11
4	CP-99,994	0.17 ± 0.04	16	**9**		0.34 ± 0.07	17
5	CP-210,053	2[a]	23				

a) Determined as the racemate.

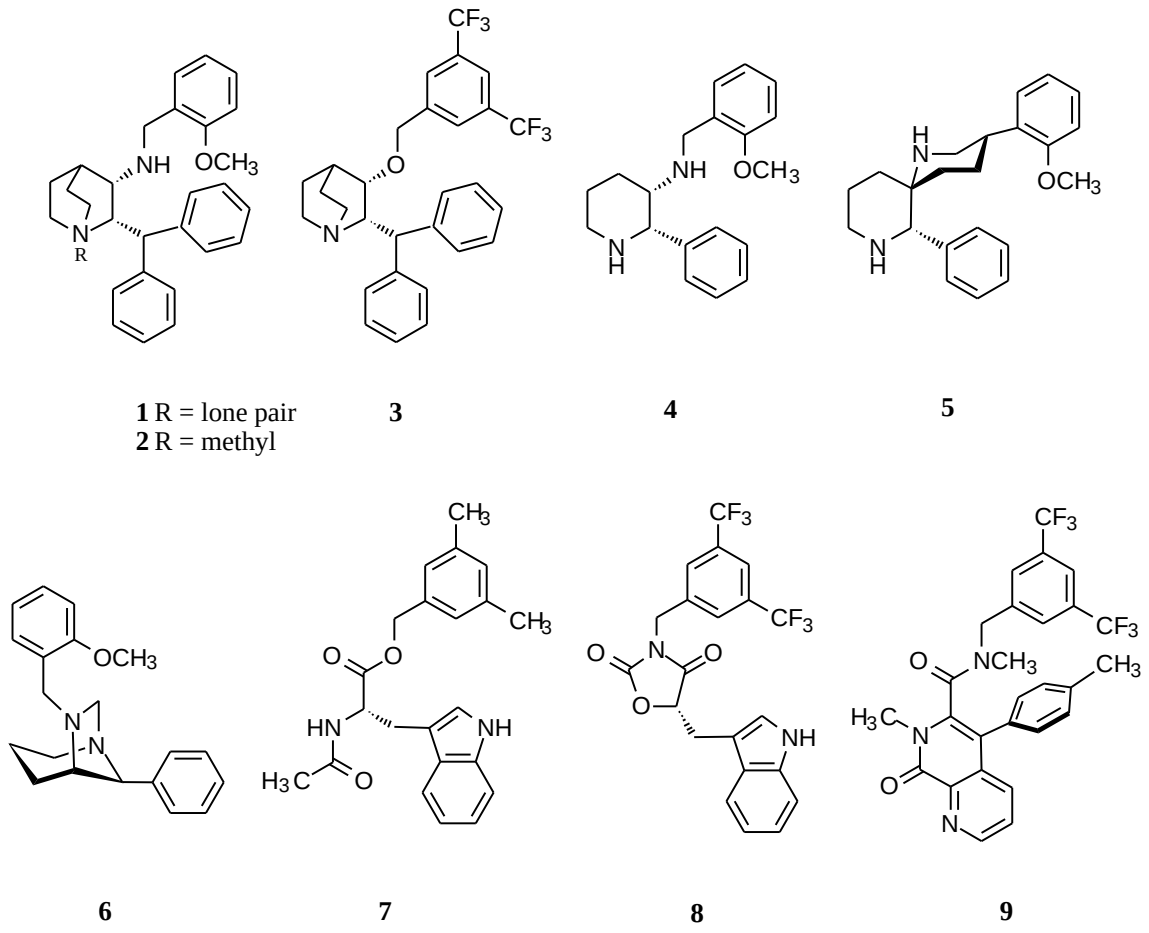
A generalized (non-peptide) NK_1 antagonist pharmacophore is shown in Fig. 1(a). This pharmacophore merely reflects the requirement of two aromatic rings for high-affinity NK_1 antagonism. The aromatic rings are kept in a proper spatial arrangement by a great variety of possible scaffolds,[5] as is also reflected in the compounds included in this study (Fig. 2). They represent four different chemical classes of NK_1 antagonists: quinuclidines (**1**-**3**), piperidines (**4**-**6**), L-tryptophan benzyl esters (**7**,**8**) and isoquinoline-urea and pyrido[3,4-*b*]pyridine carboxamide (**9**) derivatives. All compounds have high-affinity (IC_{50}<15 nM) or medium affinity (IC_{50} 15 - 150 nM in our definition) for the human NK_1 receptor (Table 1). A common part of the quinuclidine and piperidine scaffolds is exemplified by a 1,2-

disubstituted ethane fragment connecting the two pharmacophoric heteroatoms[7] as shown in Figs 1(b) and 3. This structural element can tentatively also be recognized in compounds **7** and **9** (Fig. 4). Although a different detailed mechanism of interaction with the NK_1 receptor can be expected, a similar structural role for the scaffolds is hypothesized as is, for instance, indicated by the common receptor binding sites for the L-tryptophan benzylester **7** and the benzylether quinuclidine **3**.[8] With the aid of the crystal structure data of the antagonists investigated we were able to suggest interaction geometries between these compounds and Gln-165, His-197, Phe-268, Tyr-272 and His-265 in the human NK_1 receptor.

METHODS AND MATERIALS

Crystal Structures

The crystal data and atomic coordinates were either taken from the supplementary material of the corresponding reports or from the Cambridge Structural Database (CSD).[9] The crystal data for compound **6**, CP-211,754 was kindly provided by Dr H.R. Howard.[10] Data for compounds **8** and **3** were kindly provided by Dr R.T. Lewis and Dr R.G. Ball.[11,12] In cases where atomic positions for hydrogens were absent, they were calculated using the CRYSIN module of the SYBYL package (Version 6.2).[13] A summary of selected crystallographic data for the set of antagonists analyzed is presented in Table 2. Five compounds have been cocrystallized with other small molecules (**1b**-**5**), *viz.* counter-ions, water or methanol. Compound **1** (CP-96,345) is present in two crystal forms: the free base (**1a**) and as the dimesylate salt (**1b**). The protonation state of CP-96,345 in **1b** has not been reported.[14] However, the stoichiometry of the crystal, the basicity of the nitrogen atoms, the acidity of the methanesulfonic acid together with the resulting hydrogen bond network revealed that both amine nitrogens in **1b** are protonated. The protonation state of the other compounds as reported for these crystal structures was confirmed by our analyses. Six crystal structures are encountered in centrosymmetric space groups. For the chiral compounds (**1a**, **1b**, **3**, **4** and **5**) this implies that they have been crystallized as racemic mixtures of two enantiomers. This should be kept in mind in relation to the pronounced stereoselectivity of some of these compounds.[12,15,16] Compound **9** is achiral, which, in view of the centrosymmetric space group, implies that two conformations of the same molecule, which are mirror images of one another, are present in the crystal structure.[17] For compound **8** only the crystal structure of the inactive *R* enantiomer was available to us,[11] but inversion of the



1 R = lone pair **3** **4** **5**
2 R = methyl

6 **7** **8** **9**

FIGURE 2. Structural formulas of the non-peptide NK_1 antagonists included in this study.

complete crystal structure is allowed without compromising the integrity of the X-ray diffraction experiment. A general error check on all structures was performed using PLATON.[18]

Evaluation of Interactions

In the analysis of intermolecular contacts, molecules and/or groups in van der Waals contact or at least surrounding the active stereoisomer in the crystal structures were identified using the SYBYL package.[13] Molecules were considered to be 'in van der Waals contact' or 'surrounding' if in any pair of non-hydrogen atoms, not belonging to the same molecule, the atoms were no more than 5 Å apart. Intermolecular contacts for functional groups that are

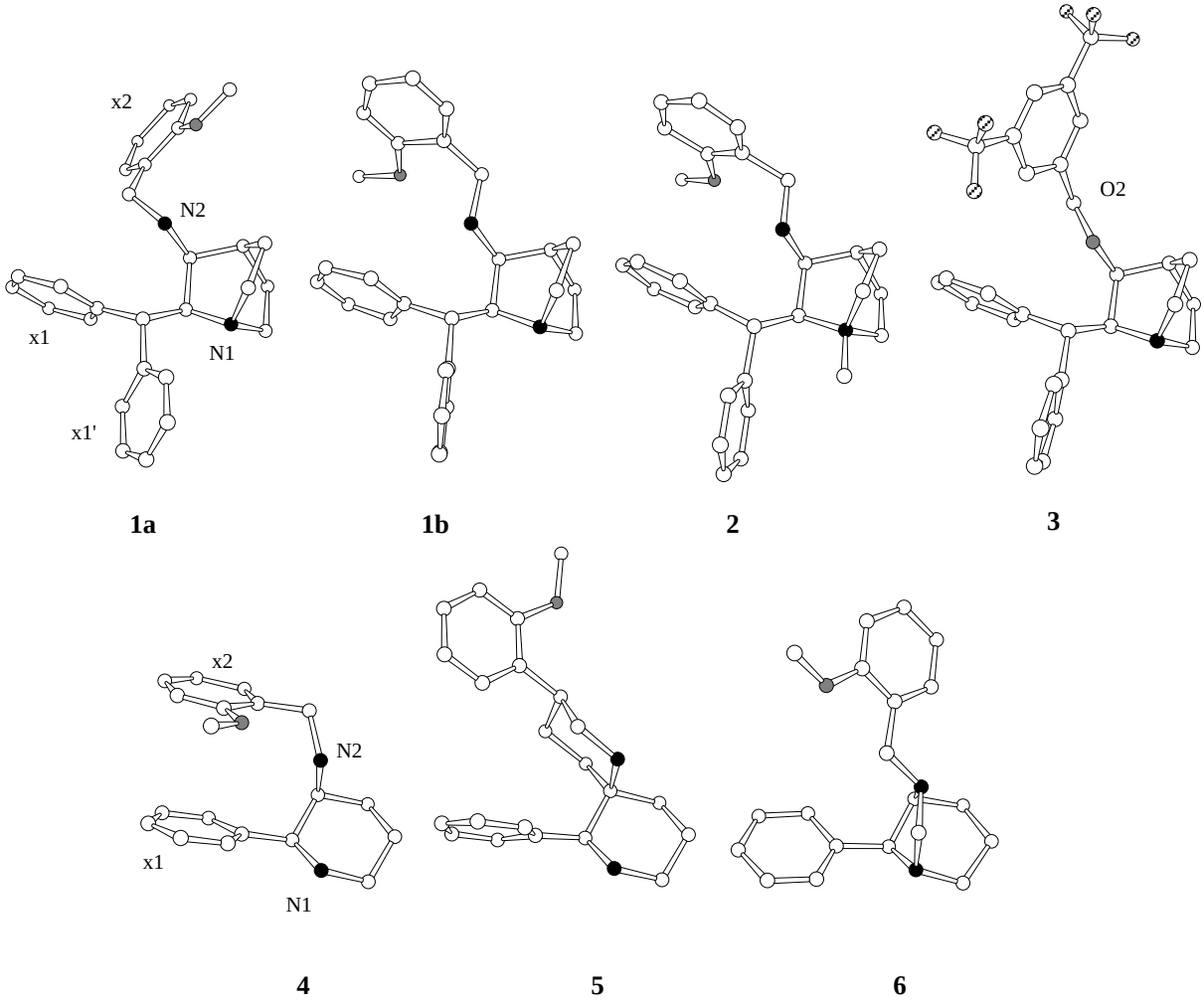
considered to be important for the activity of the compounds were classified and geometrically evaluated in terms of documented intermolecular interactions (see e.g. Burley and Petsko[19]). For details on the determination of interaction geometries see the Experimental section.

Hydrogen bond analyses of all structures in their established protonation state and with their crystallographically determined hydrogen positions, were performed with PLATON.[18] For the evaluation of possible C-H··X interactions, X being N,O or Cl and I (usually as anions),[20] hydrogen bond criteria were applied (see the Experimental section). For these analyses corrected hydrogen positions in the structures have been calculated using the CRYSIN module of the SYBYL package. These new hydrogen positions relate to the normalisation of bond lengths involving hydrogen atoms; the values determined by X-ray crystallography are systematically shorter by about 0.1 Å than their actual values, which can

TABLE 2. Selected crystallographic data for the NK_1 antagonists analyzed

No.	Name	Formula	Space group	*R*-factor[a]	CSD Refcode and ref.
1a	CP-96,345 free base	$C_{28}H_{32}N_2O$	P-1	0.040	YAFJOE, 14
1b	CP-96,345 dimesylate salt	$C_{28}H_{32}N_2O{\cdot}2CH_3SO_3H$	P-1	0.123[b]	YAFJUK, 14
2	*N*-methyl analogue of CP-96,345[c]	$[C_{29}H_{35}N_2O^+][I^-]$	$P2_1$	0.072	LEWCUL, 17
3	L-709,210 hydrochloride[d]	$C_{29}H_{27}NOF_6{\cdot}HCl$	P-1	0.096	12[e]
4	CP-99,994 dihydrochloride	$C_{19}H_{24}N_2O{\cdot}2HCl{\cdot}0.05CH_3OH$ [f]	$P2_1/n$	0.055	LACPOU, 16
5	CP-210,053 dihydrochloride	$C_{22}H_{28}N_2O{\cdot}2HCl{\cdot}H_20$ [g]	Pbca	0.060	YICLEB, 23
6	CP-211,754	$C_{20}H_{24}N_2O$	$P2_1$	0.035	10[e]
7	L-708,568	$C_{22}H_{24}N_2O_3$	$P2_1$	0.045	26
8	L-732,244[d]	$C_{21}H_{14}N_2O_3F_6$	$P2_12_12_1$	0.057	11[e]
9		$C_{27}H_{21}N_3O_2F_6$	$P2_1/n$	0.096	17

a) *R*-factor $=\sum||F_o|-|F_c||/\sum|F_o|$, in which $|F_o|$ and $|F_c|$ are the moduli of the observed and calculated structure factors, respectively. b) The authors report that this relatively high *R*-factor is likely to be due to the low quality of the crystal related to its sensitivity to the atmosphere. c) The antagonist carries a permanent positive charge; it has been crystallized as the monoiodide salt. d) The trifluoromethyl groups were rotationally disordered and modelled over two sets of staggered conformations. e) This structure will be available in future releases of the Cambridge Structural Database (CSD). f) The methanol was found to be disordered with an occupancy of 0.05. Hydrogen positions on methanol were not located. g) Hydrogen positions on the water molecule were disordered.



1a **1b** **2** **3**

4 **5** **6**

FIGURE 3. Crystal structure conformations of the quinuclidine (upper row) and piperidine (bottom row) antagonists. Atom labelling of pharmacophoric groups and selected atoms is indicated for representative compounds in each row. x1, x1' and x2 are used to indicate both the aromatic ring itself as well as the ring's centroid.

be determined by neutron diffraction.[21] Additionally, all C-H groups not considered as being involved in C-H··X interactions, but with hydrogen atoms within the van der Waals radius of any possible acceptor atom were identified. This facilitated the localization of possible patterns of C-H··X contacts, which might be more informative than restricting oneself to the interactions revealed by strictly applying the hydrogen bond criteria.[22]

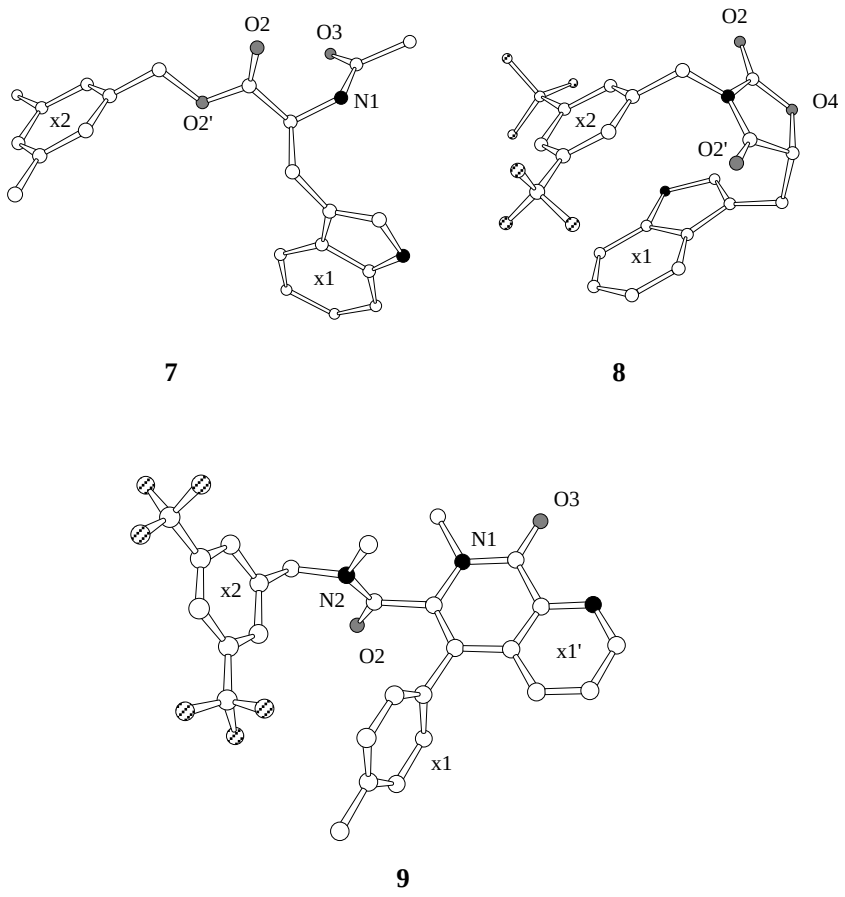


FIGURE 4. Crystal structure conformations of the L-tryptophan benzyl ester (upper row) and pyrido[3,4-*b*]pyridine carboxamide (bottom row) antagonists. For compound **8** the crystal structure conformation of the inactive *R* enantiomer is depicted. For compound **9** only one of the two mirror image conformations present in the structure is shown. Atom labelling of pharmacophoric groups is indicated for representative compounds in each row.

In the course of the study the need was felt to include molecular models of the NMR conformation[23] of compound **5** and the nitrogen N2 inverted configuration of compound **6** for comparison to the crystal structure conformations. A combination of molecular dynamics simulations followed by semi-empirical geometry optimizations using the AM1 Hamiltonian[24] in the MOPAC93 program[25] provided these models. See the Experimental section for details. Calculations were performed on Silicon Graphics INDY and Crimson Elan workstations.

RESULTS

The crystal structure conformations of the quinuclidine and piperidine antagonists are depicted in Fig. 3. The crystal structure conformations of compounds **7**,[26] **8** and **9** are depicted in Fig. 4. The assignment of the aromatic ring centroids x1 and x1' in structures **1a**, **1b**, **2**,[27] **3** and **9** (Figs 3 and 4) is based both on the mutual correspondence of the x1-x2 and x1'-x2 distances in these structures (data not shown), and the correspondence of the x1-x2 distance in CP-96,345 (**1a**, **1b**) and CP-99,994 (**4**) in molecular modelling studies as described by Desai et al.[16] Intermolecular interactions encountered in the crystal structures are (charge-assisted) hydrogen bonds and N^+(charged nitrogen) - aromatic interactions for N1, C-H··X interactions (X being O, chloride or iodide) for carbon atoms adjacent to a positively charged nitrogen atom, and aromatic-aromatic stackings and hydrophobic interactions for the aromatic groups in the molecules. Below we describe these interactions and their interaction geometries in more detail, with an emphasis on the interactions of the pharmacophoric groups.

Quinuclidine Antagonists

Whenever the quinuclidine nitrogen N1 is positively charged (**1b**, **2** and **3**), there are always two anions present at similar positions in space. These positions are indicated as I and II in Fig. 5 and Table 3. These anion positions may mimic the positions of hydrogen bond acceptor groups or anionic groups in the NK_1 receptor. The anion positions as reported in Table 3 were given relative to the quinuclidine rings. When the positive charge of N1 is due to protonation, the anion in position I is involved in a charge-assisted hydrogen bond with N1. In structure **2** the iodide anion in position I is involved in a C-H··I^- interaction with the methyl group attached to N1. In position II the anion positions are stabilized by C-H··X interactions arising from either C2-H2 (**2**) or from both C2-H2 and C7-H7a (**1b**, **3**) (Fig. 5). The similarity in anion positions in the various crystal structures is reflected in the dihedral angles τ(N1-C2-C3-Ac) for positions I and II as given in Table 3. In the only structure with a neutral quinuclidine nitrogen (**1a**), positions I and II are occupied by phenyl rings. These positions therefore probably correspond to general packing sites.

In addition to the anions, the quinuclidine rings carrying a positively charged N1 nitrogen are surrounded by phenyl rings stabilized by N^+-aromatic interactions.[28,29] These

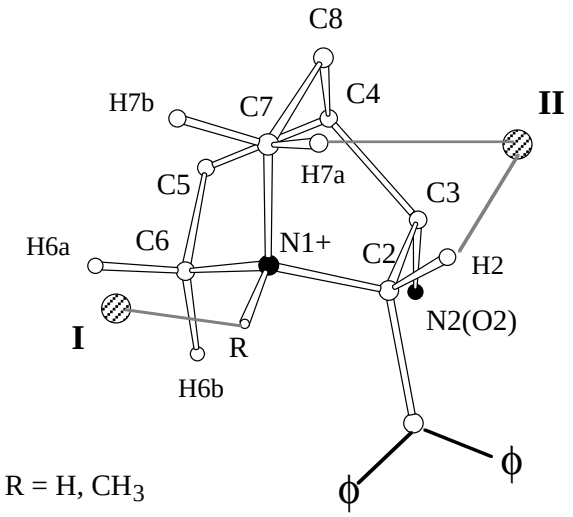


FIGURE 5. Common anionic interaction sites for quinuclidine NK_1 antagonists. The crystal structure of L-709,210 (**3**) serves as an example.

TABLE 3. Common intermolecular hydrogen bond acceptor and/or C-H··X contact acceptor groups surrounding quinuclidine antagonists

Position I						
No.	Donor	Acceptor[a]	d(N1··Ac) (Å)[b]	<(C2-N1··Ac) (˚)	τ(C2-N1-R-Ac)[c] (˚)	τ(N1-C2-C3-Ac) (˚)
1b	N1-H	O(5) mesylate	2.686(19)	130.7	-176.5	-7.2
	C6-H6a	O(7) mesylate	3.91(2)	153.6	-155.5	-18.1
2	C1-H[d]	I(1) iodide	4.652	165.5	-172.7	-6.8
3	N1-H	Cl(1) chloride	3.113(5)	136.0	152.0	8.4
Position II						
No.	Donor	Acceptor	d(N1··Ac) (Å)	<(N1-D··Ac) (˚)	τ(N1-D-H-Ac) (˚)	τ(N1-C2-C3-Ac) (˚)
1b	C2-H2	O(6)' mesylate	3.97(2)	101.1	-15.8	108.7
	C7-H7a	O(5)' mesylate	4.21(2)	116.4	167.5	76.1
2	C2-H2	I(1)' iodide	4.845	111.0	-109.3	113.2
3	C2-H2	Cl(1)' chloride	4.107(5)	91.8	-8.8	98.9
	C7-H7a	Cl(1)' chloride	4.107(5)	99.4	68.6	98.9

Atom labelling as given in Fig. 5. Ac and D denote an acceptor group and a donor carbon atom, respectively. In the Table the anion positions have been defined relative to the antagonist scaffolds. a) The acceptor atoms have been labelled in accordance with the original publication. A prime indicates a position related by crystallographic symmetry to the acceptor atom mentioned earlier in the Table. b) Values in parentheses are estimated standard deviations (e.s.d.'s) in the last digit. For compound **2** no e.s.d.'s were available. c) R is either the N1 attached hydrogen atom, or the C1 carbon atom of the methyl group attached to N1 in case of compound **2**. d) C1 is the carbon atom of the methyl group attached to N1.

phenyl rings are all positioned on one hemisphere of the quinuclidine ring, as can be seen in the superposition in Fig. 6. Some positions are exclusively occupied by aromatic rings while others are, occasionally, also occupied by anions in other structures. Interaction geometries are given in Table 4. Position *a*, above C7 in Fig. 6, is occupied by a phenyl ring in three structures (**1b**, **2** and **3**) and the ring is tilted from a perpendicular position with respect to the vector N1-C7. However, this position is also occupied by a phenyl ring in the crystal structure of the neutral quinuclidine (**1a**). Position *d* is occupied by a phenyl ring in structure

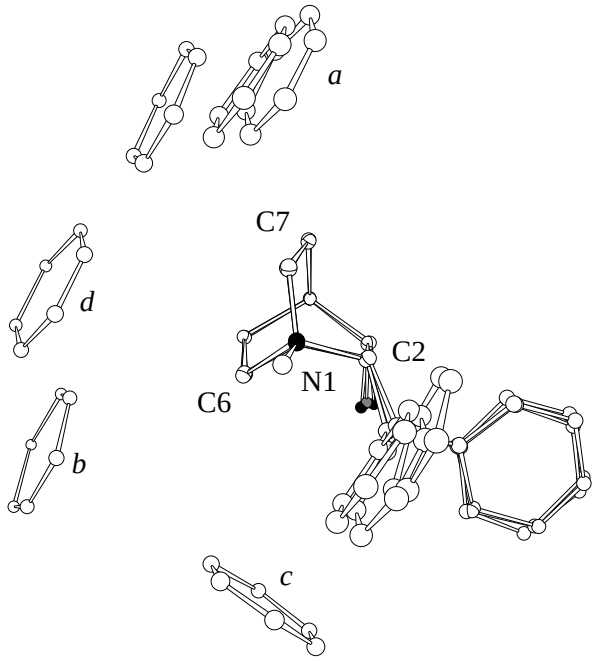


FIGURE 6. N^+-aromatic interaction geometries as observed around the positively charged quinuclidine rings. The quinuclidine rings have been superimposed. The substituted benzyl groups of the antagonists have been omitted for clarity.

3 and by a mesylate anion in structure **1b**, both making van der Waals contact with the quinuclidine ring. The phenyl ring position designated *c* (structure **2**) is stabilized by both N+-aromatic interactions and possible C-H··π interactions[30] donated by the benzhydryl group.

In contrast to N1, the pharmacophoric center N2(O2) is rarely involved in interactions in the crystal structures of the quinuclidine antagonists; it forms a charge-assisted hydrogen bond with a mesylate anion in structure **1b** and geometrically non-ideal intramolecular hydrogen bonds with the 2-methoxy group of x2 in structures **1a** and **1b** (Fig. 3). The N2(O2) atom of the other quinuclidine structures and the 2-methoxy oxygen atoms in **1a**, **1b** and **2** are not involved in any specific intermolecular interactions.

TABLE 4. Observed N^+-aromatic interaction geometries for quinuclidine antagonists in their crystal structures

No.	C(N1+)	Position in Fig. 6	Type[a]	d(N1··x) (Å)	d(C(N1+)··x) (Å)	θ (°)	ϕ (°)	τ(C2-N1-C(N1+)··x) (°)
1b	C7	*a*	2	5.618	4.258	20.5	46.4	-123.4
2	C7	*a*	2	5.417	4.057	22.3	36.7	154.4
	C6	*b*	2	5.993	4.543	16.8	55.8	162.0
	C6	*c*	1	5.292	4.749	61.3	29.2	72.5
	C1	*c*	1	5.292	5.032	71.9	29.2	-75.4
3	C7	*a*	2	5.650	4.251	18.3	53.6	-101.3
	C7	*d*	1	5.268	5.015	72.1	35.7	142.3
	C6	*d*	1	5.268	4.273	41.8	35.7	-149.9

The geometrical descriptors as applied for N^+-aromatic interactions for the quinuclidine antagonists are defined as follows[29]: θ is the angle <(C(N1+)-N1··x) (x=centroid), ϕ is the angle between the normal of the plane of the phenyl ring and the elongation of the vector from N1 to x. C(N1+) indicates the carbon atom adjacent to N1+. The additional dihedral angle τ(C2-N1-C(N1+)··x) is given to reconstruct the position of the phenyl ring centroid in space. Note that the orientation of the phenyl ring is not fully determined by the parameters tabulated. a) The type of interaction geometry is given in accordance with the preferential geometries given by Verdonk et al.[29]

In structure **3** the disubstituted benzylether phenylring x2 is involved in an internal edge-to-face aromatic stack with x1 (Fig. 3). The x2 ring is at the same time positioned parallel (stacked) in between two phenyl groups of neighbouring molecules (Fig. 7). One of these phenyl groups is a 3,5-disubstituted phenyl ring with which the parallel 'stacking' geometry is accomplished by the trifluoromethyl groups attached to both phenyl rings; these groups are located directly above the other ring centroid. No specific interactions of the benzylamino phenyl group x2 in structures of **1a**, **1b** and **2** were observed.

The benzhydryl groups of CP-96,345 in both crystal structures **1a** and **1b**, form symmetric dimers with the benzhydryl group of neighbouring molecules which are related to

the first one by inversion (Figs 8(a) and (b), respectively). In structure **1a** this geometry is stabilized by two edge-to-face aromatic stacks and in **1b** by one face-to-face and two edge-to-face aromatic stacking interactions. Additionally, in structure **1a** the benzhydryl group of the neutral quinuclidine coordinates a phenyl ring in a cooperative fashion (Fig. 8(c)). These interaction geometries might reflect the possible interactions between the benzhydryl group of **1**-**3** and the aromatic side chains of His-197, Phe-268 and Tyr-272 in the human NK_1 receptor.[3] The benzhydryl groups of compounds **2** and **3** only coordinate, by hydrophobic interactions, a methoxy methyl group and the hydrophobic side of a symmetry related quinuclidine ring, respectively, similar to the interaction geometry in Fig. 8(c).

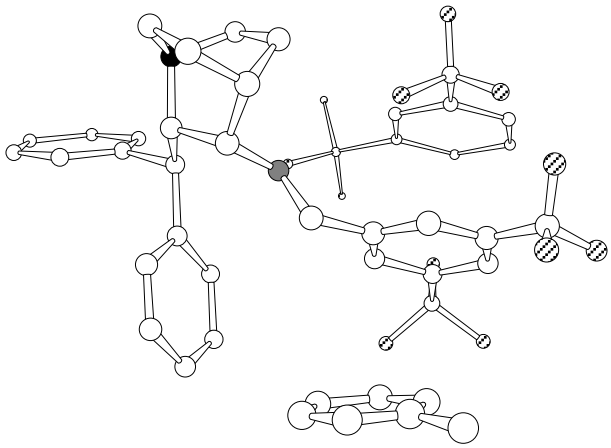
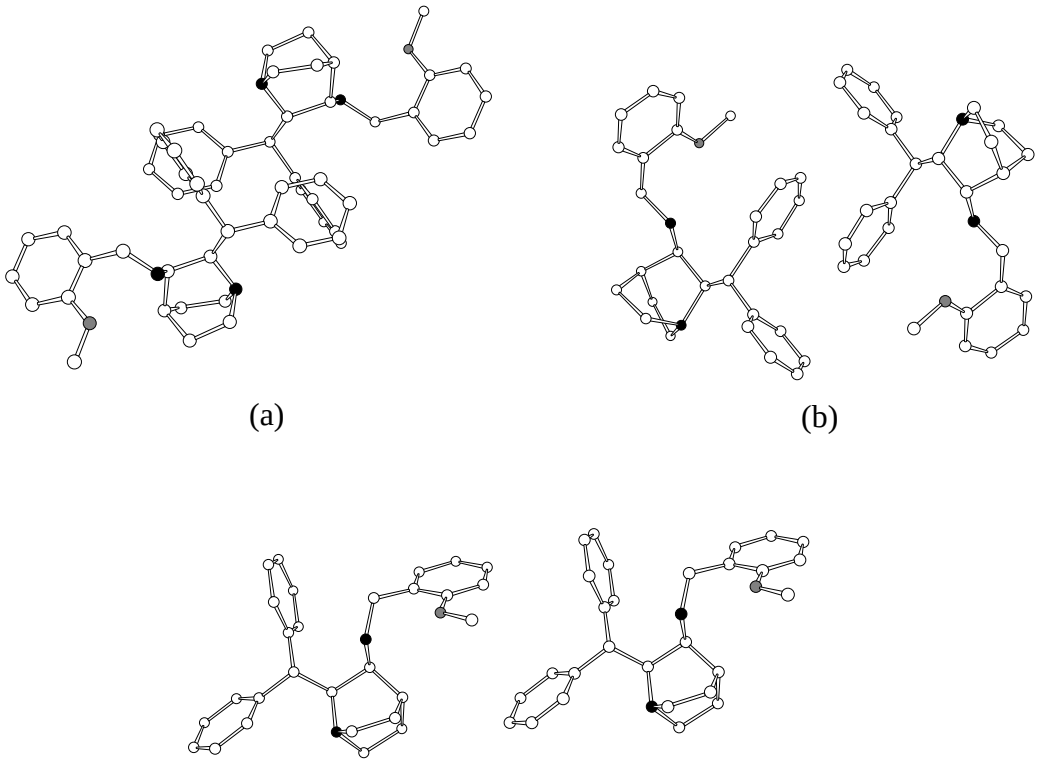


FIGURE 7. Parallel stacking of the benzylether phenyl ring of L-709,210 (**3**) by two neighbouring phenyl rings. In case of the nearby disubstituted phenyl ring (top), the stacking is mutually accomplished by the trifluoromethyl groups.

Piperidine Antagonists

In two out of three piperidine antagonist structures, CP-99,994 (**4**) and CP-210,053 (**5**), both N1 and N2 are protonated, resulting in complete surrounding of their scaffold parts by hydrogen bond acceptor groups, i.e. either chloride anions or, in one case (**5**), a water molecule. The positions of these groups are similar (designated as A, B, C and D) in the two structures and are shown in Fig. 9 and indicated in Table 5. The chloride anion in position C is, in contrast to the other hydrogen bond acceptor groups, coordinated by two C-H··Cl^- interactions. In the crystal structure of compound **4** two phenyl rings are partially located within the 5 Å coordination sphere around N1, close to C6 and can probably be regarded as being involved in N^+- aromatic interactions, even if no strict geometrical criteria can be

applied; the N1-centroid distances are 5.24 and 5.47 Å. In the neutral piperidine antagonist **6**, no specific interactions for N1 and N2 were observed.



(a) (b)

(c)

FIGURE 8. Aromatic-aromatic interaction geometries of the benzhydryl group of CP-96,345 with surrounding phenyl rings. (a) CP-96,345 free base (**1a**), (b) CP-96,345 dimesylate salt (**1b**), (c) Additional hydrophobic coordination of a phenyl ring by the benzhydryl group of **1a**.

A major difference in the intermolecular interactions of the doubly protonated piperidines **4** and **5** is observed in the interactions of their aromatic rings. The phenyl rings in the crystal structure conformation of CP-99,994 (**4**) are involved in an internal face-to-face aromatic stack, which is stabilized by an intramolecular hydrogen bond between nitrogen N2 and the 2-methoxy group. In addition, both phenyl ring x1 and x2 are involved in an intermolecular edge-to-face aromatic stack (Fig. 10(a)). In contrast, the phenyl rings of

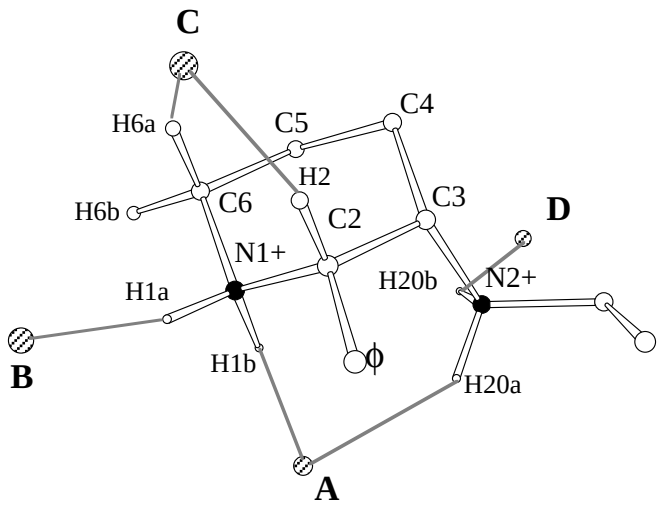


FIGURE 9. Common hydrogen bond acceptor and/or C-H··X contact acceptor group positions for piperidine NK_1 antagonists. The structure of CP-99,994 (**4**) serves as illustration.

TABLE 5. Common hydrogen bond acceptor and/or C-H··X contact acceptor groups surrounding the piperidine antagonists **4** and **5**[#]

Position A							
No	Donor	Acceptor[a]	d(D··Ac) (Å)	<(C2-N1·Ac) (˚)	<(C3-N2·Ac) (˚)	τ(C3-C2 - N1-Ac) (˚)	τ(C2-C3 -N2- Ac) (˚)
4	N1-H1b	chloride Cl(2)	3.104(5)	110.4	-	61.8	-
	N2-H20a	chloride Cl(2)	3.261(5)	-	110.3	-	38.5
5	N1-H1b	chloride Cl(1)	3.126(7)	111.4	-	54.9	-
	N2-H20a	chloride Cl(1)	3.161(6)	-	113.3	-	38.0
Position B							
No	Donor	Acceptor	d(N1··Ac) (Å)	<(C2-N1·Ac) (˚)	τ(C3-C2-N1-Ac) (˚)		
4	N1-H1a	chloride Cl(1)'	3.063(5)	114.2	-168.9		
5	N1-H1a	water O(26)	2.762(10)	93.5	-164.9		
Position C							
No	Donor	Acceptor	d(N1··Ac) (Å)	<(N1-C2·Ac) (˚)	<(N1-C6·Ac) (˚)	τ(C3-C2-N1-Ac) (˚)	
4	C2-H2	chloride Cl(2)'	4.131(5)	96.9	81.0	-119.8	
5	C6-H6a	chloride Cl(2)'	3.787(8)	83.8	88.5	-116.3	

[#] Table is continued on the next page.

Table 5 *Continued*

			Position D		
No.	Donor	Acceptor	d(N2··Ac) (Å)	<(C3-N2·Ac) (°)	τ(C2-C3-N2-Ac) (°)
4	N2-H20b	chloride Cl(1)''	3.174(4)	115.8	129.8
5	N2-H20b	chloride Cl(2)''	3.054(7)	119.3	149.7

The Table has been constructed in such a way that the hydrogen bond acceptor positions are given relative to the antagonist structures. a) Acceptor atoms are labelled in accordance with the atom names in the original publication. Symmetry related atoms are indicated by a prime or double prime.

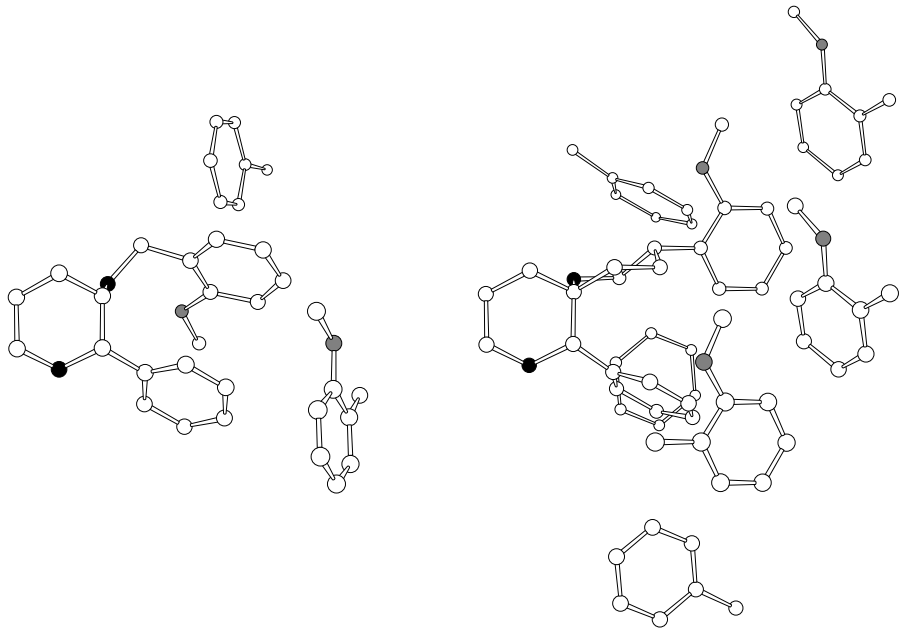
CP-210,053 (**5**), which are involved in an internal edge-to-face stack, are each involved in *three* edge-to-face aromatic stacks with surrounding phenyl rings (Fig. 10(b)). A three-fold edge-to-face aromatic stacking situation is also observed for phenyl ring x2 of the neutral piperidine antagonist **6**, whereas phenyl ring x1 is involved in one edge-to-face aromatic stack (data not shown).

Molecular dynamics simulations on the crystal structure conformations of the conformationally constrained piperidine antagonists **5** and **6** (data not shown) revealed that the N2-H and N2 lone-pair directions are greatly restricted and these compounds might therefore be used in determining the bioactive conformation of the piperidine antagonists. In addition to the crystal structure conformations of **4**, **5** and **6** models were constructed for the solution structure of **5** in $CDCl_3$ as determined by NMR[23] and the N2-inverted configuration of **6**. These five conformations are probably all important energetically accessible conformations and all correspond to staggered conformations around C3-N2 (Table 6, first data column). Upon superposition of the piperidine rings it was noted that there are three regions in space where functional groups interacting with N2 can be located. This is shown in Fig. 11. Two of these regions are exemplified by the chloride anion positions A and D in structures **4** and **5**. The third region is indicated by e.g. the N2 (axial) lone pair in the solution conformation of **5** (Table 6 and Fig. 11).

L-Tryptophan Benzyl Ester Antagonists

In the crystal structures of the L-tryptophan benzyl ester analogs **7** and **8** only few heteroatoms are involved in hydrogen bonds, perhaps because of the low number of

hydrogen bond donors in the molecules. Comparison of the crystal structure conformations of **7** and the conformationally constrained analogue **8** (Fig. 4) on the one hand suggests that the crystal structure conformation of **7** is not the bioactive conformation: the distance between the centroids of the aromatic rings x1 and x2 in **7** is more than twice this distance in the structure of **8**. On the other hand, the distance between x1 and x2 in the crystal structure of **7** (i.e. 8.15 Å) corresponds quite well with the x1'-x2 distances in the crystal structures of **1b**-**3**, which range from 8.38 Å in **3** to 8.76 Å in **1b**. The indole ring of **7** might therefore also be considered able to occupy the same region as the x1' group of the quinuclidine antagonists,



(a) (b)

FIGURE 10. Edge-to-face aromatic - aromatic interaction geometries for the phenyl and benzylamino groups in the crystal structures of the doubly protonated piperidine antagonists (a) CP-99,994 (**4**) and (b) CP-210,053 (**5**).

when bound to the NK_1 receptor. Anyhow, both in structures **7** and **8** the 3,5-disubstituted phenyl rings (x2) are involved in intermolecular edge-to-face aromatic stacks as has been reported previously.[11,26] In structure **8** this 'stacking' involves the rotationally disordered trifluoromethyl groups.

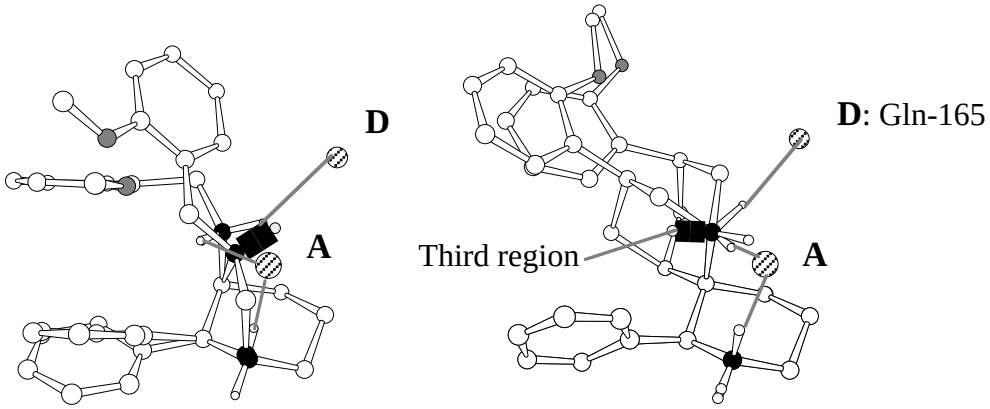


FIGURE 11. Regions available to groups interacting with N2 in the piperidine antagonists **4**, **5** and **6**. Structures have been superimposed based on their piperidine rings. The structures of CP-99,994 (**4**) and **6** are shown on the left. The N2-inverted configuration of **6** is not shown. The crystal and solution conformations of CP-210,053 (**5**) are shown on the right with the substituted phenyl ring in an equatorial and axial position, respectively. Lone pairs are depicted in black. Hydrogen bond acceptor positions are labelled in accordance with Fig. 9.

TABLE 6. Staggered conformations around C3-N2 and interaction sites for N2 for the crystal structure as well as the calculated conformations of the piperidine antagonists

No.	Conformation	τ(C2-C3-N2-C9) (˚)	τ(C2-C3-N2-A) (˚)	τ(C2-C3-N2-D) (˚)	τ(C2-C3-N2-lp) (˚)
4	crystal structure	-135.3	38.5	129.8	-
5	crystal structure	-85.2	38.0	149.7	-
	solution (calculated)	-170.0	40.8	-	-45.6
6	crystal structure	-100.1	(22.5)[a]	150.2 or 164.0[b]	-
	N2-inverted (calculated)	156.3	(28.5)[a]	-	-86.0

Atom labelling as given in Fig. 9, lp denotes a lone pair. a) In these structures position A is occupied by a carbon atom. b) The values are determined to anion position D of crystal structures **4** and **5**, respectively.

Pyrido[3,4-*b*]pyridine Carboxamide Antagonist

No hydrogen bonds have been observed in the crystal structure of **9**. It seems, however, that the lack of hydrogen bond donors in the crystal is compensated for by short C-H··O contacts with e.g. O2 and O3 (Fig. 4). The pyrido[3,4-*b*]pyridine rings of two symmetry

related (in this case mirror image conformations) molecules are stacked in a perfect (anti)parallel manner. In addition, the aromatic rings x1 and x1', which topologically resemble a 'frozen' benzhydryl group, cooperatively coordinate a trifluoromethyl group.

Superposition of Quinuclidine and Piperidine Antagonists

Desai et al.[16] postulated the pharmacophoric dihedral angle x1-C2-C3-N2(O2) as a descriptor of the structural similarity of quinuclidine and piperidine antagonists. To investigate whether this structural similarity is reflected in the intermolecular interactions in the crystal structures, the crystal structure conformations of the quinuclidines **1b**, **2** and **3** were superimposed with the crystal structure conformation of CP-99,994 (**4**) based on a least-squares fit of the corresponding atom pairs x1, C2, C3 and N2(O2). The resulting RMSd values were 0.405, 0.429 and 0.422 Å, respectively. As a typical example, the resulting superposition of CP-96,345 (**1b**) with CP-99,994 (**4**) is depicted in Fig. 12. The respective N2 and x1

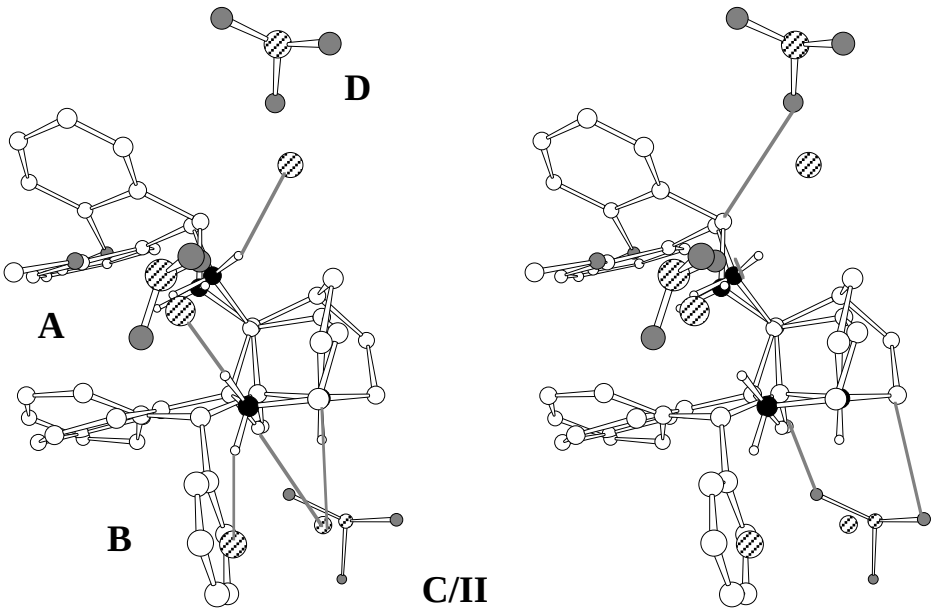


FIGURE 12. Stereoview of the superposition of the crystal structure conformations of CP-99,994 (**4**) and CP-96,345 (**1b**) based on x1, C2, C3 and N2, RMSd=0.405 Å. Hydrogen bond acceptor positions are labelled in accordance with Figs 5 and 9. Dashed lines indicate hydrogen bonds or C-H··X interactions in **4** (left) and **1b** (right). The mesylate methyl groups have been omitted for clarity.

positions match fairly well, while the N1 positions are clearly distinct. Also, the chloride anion in position C that is common for the piperidine antagonists **4** and **5** was positioned at or nearby the anions at position II of the quinuclidine antagonists. In contrast, the piperidine anion position B is located close to the x1' centroid of the benzhydryl group of the quinuclidines. One of the two aromatic rings that can probably be considered to be involved in N^+-aromatic interactions with the positively charged head group of the piperidine ring of CP-99,994 (**4**) (*vide infra*) is in close proximity of the positions of both phenyl ring *b* in structure **2** and phenyl ring *d* in structure **3** that are depicted in Fig. 6; the intercentroid distances in the superpositions are 2.9 and 2.2 Å, respectively.

In the exemplified case (Fig. 12) of the superposition of CP-96,345 (**1b**) and CP-99,994 (**4**), which are both doubly protonated, two additional anion positions are found at similar sites in space: positions A and D. The mesylate anion in position D is bonded by C-H··O interactions.

DISCUSSION

Crystal structures of nine non-peptide tachykinin NK_1 antagonists have been analyzed for the intermolecular interactions of their pharmacophoric groups with neighbouring molecules in the crystal. Several interaction geometries have been identified which are shown to be consistent with both structure-activity relationships and reported receptor interactions for the compounds analyzed. A schematic representation of reported ligand - receptor interactions combined with our results for the quinuclidine and piperidine antagonists **1**-**6** is shown in Fig. 13. The interactions encountered in the crystal structures are charge-assisted hydrogen bonds (salt-bridges), C-H··X interactions (X being O, chloride or iodide), N^+-aromatic interactions, aromatic-aromatic interactions and hydrophobic interactions.

The quinuclidine antagonists reveal two common hydrogen bond acceptor sites, whenever N1 is positively charged (Fig. 5). Structure-activity studies show that the N1 nitrogen is likely to be positively charged in the active site,[14,27] either by protonation or alkylation as in compound **2**. The rigidity of the quinuclidine framework ensures that no major differences exist between the quinuclidine ring conformation in the crystal structure and in the bioactive state. The same is therefore true for the hydrogen bond acceptor sites. Since the common hydrogen bond acceptor (anion) sites occur both with single-atom anions as well as with the more complex mesylate anions, it stresses the importance of these interactions for a quinuclidine ring containing a positively charged nitrogen. Although the

general importance of these hydrogen bond acceptor sites is yet to be confirmed by e.g. crystal structure statistics, it is likely that these interaction sites are relevant candidates for hydrogen bond acceptor groups in the NK_1 receptor. These are indicated by 'HAcc' in Fig. 13.

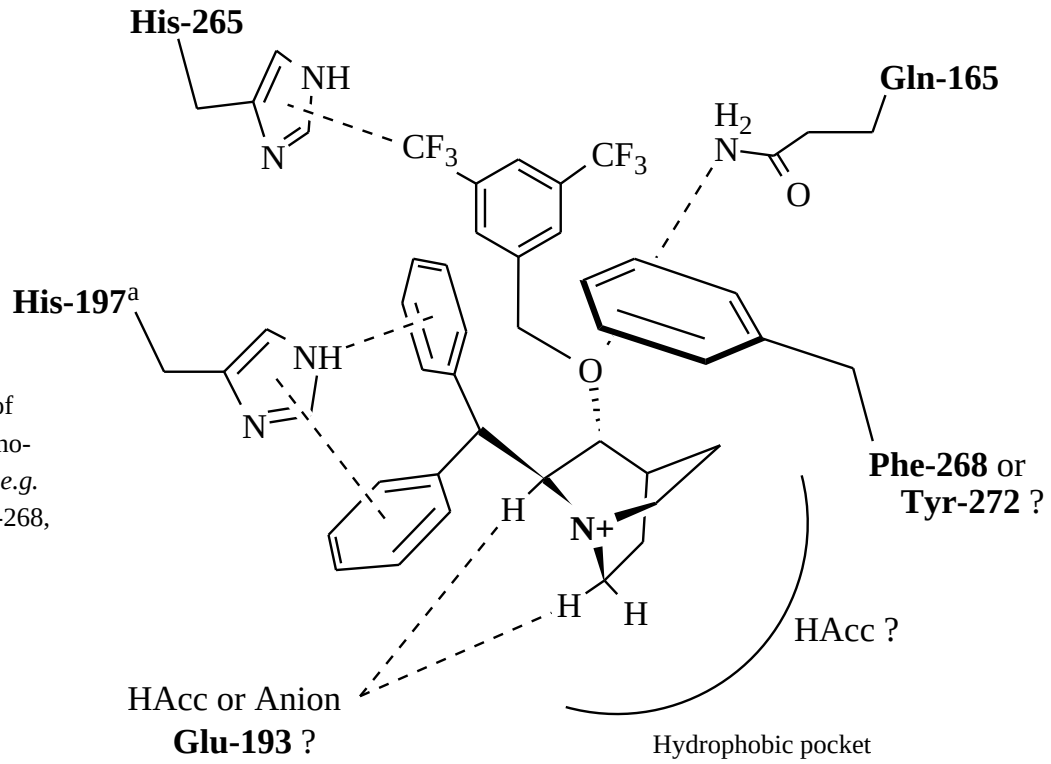


FIGURE 13. Schematic representation of reported ligand - receptor interactions combined with our results for the quinuclidine and piperidine antagonists **1**-**6**. 'HAcc' stands for a hydrogen bond acceptor group in the NK_1 receptor. The dashed lines between His-265 and the 3,5-disubstituted phenyl ring and between the centroid of the imidazole ring of His-197 and one of the benzhydryl phenyl rings indicate reported interactions[8,32,35] without inferring the exact nature of the interaction.

The only anionic residue that is a candidate for occupying one of these hydrogen bond acceptor sites in the quinuclidine binding site in the human NK_1 receptor is Glu-193 that is reported to perform an indirect role in antagonist binding,[31] however. It should be noted that electroneutrality in the binding of a positively charged quinuclidine antagonist to the receptor can also be realized if a free anion would be involved in binding.

Aromatic amino-acid residues (including histidines) play an important role in the non-peptide antagonist binding site of the NK_1 receptor.[3] Each of the N^+-aromatic interaction sites around the protonated quinuclidine ring may therefore be occupied by aromatic amino-acid residues in binding to the NK_1 receptor (Indicated as a putative hydrophobic pocket in Fig. 13). The presence of these aromatic rings is in agreement with the hydrophobic character of the other ring structures that can replace the quinuclidine ring as a scaffold in NK_1 antagonists.[5] The preferred presence of a hydrogen bond acceptor group (anion position I, Fig. 5) close to these aromatic rings might additionally be tested in NK_1 antagonist design. One of the aromatic interaction sites (position *c* in Fig. 6 and tentatively indicated as Phe-268 or Tyr-272 in Fig. 13) is especially interesting, because of a cooperative effect of N^+-aromatic and multiple C-H··π interactions with the quinuclidine and benzhydryl groups, respectively. This interaction site might therefore be specific for the quinuclidine antagonists.

The benzhydryl group of the quinuclidine antagonists is thought to be bound by His-197 of the human NK_1 receptor via amino-aromatic interactions[32] (Fig. 13) or by a cluster of aromatic amino-acid residues (His-197, Phe-268 and Tyr-272).[3] Possible interaction geometries for a benzhydryl group with a cluster of aromatic rings were suggested by the aromatic stacking geometries in the two crystal structures of CP-96,345 (**1a** and **1b**), as can be seen in Fig. 8. Although these interaction geometries only present energetical and geometrical possibilities, they might be used in making binding site models of the human NK_1 receptor.

CP-99,994 (**4**) has been developed based on CP-96,345 (**1**). The inner phenyl ring of the benzhydryl group of CP-96,345 (**1**) with respect to the benzylamino nitrogen is thought to be mimicked by the C2 phenyl ring (x1) in CP-99,994 (**4**). The rationale for the correspondence in NK_1 affinity of both compounds has been given by Desai et al.[16] in terms of the dihedral angle x1-C2-C3-N2, which is similar ((+)-*synclinal*) for the quinuclidine and the piperidine antagonists in the crystal structures and in molecular mechanics calculations.[16] A superposition of the crystal structure conformations of quinuclidine and piperidine antagonists based on the elements x1, C2, C3 and N2/O2 resulted in a mismatch of N1, but in corresponding positions for N2 and x1 and hydrogen bond acceptor groups (anions) in C and II (Fig. 12). Also, a corresponding position for aromatic rings involved in N^+-aromatic interactions in two quinuclidine crystal structures (**2** and **3**) and CP-99,994 (**4**) was observed. The scaffold parts of the quinuclidine and piperidine antagonists thus show similar interactions with their environment, i.e. hydrogen bond acceptor position C/II (Figs 5 and 9)

and a N^+-aromatic interaction. For the diprotonated quinuclidine **1b** and the diprotonated antagonists **4** and **5** additional interacting hydrogen bond acceptor groups loosely coincide in space. Although the N1 monoprotonated rather than the diprotonated state of the quinuclidine and piperidine antagonists is probably the bioactive state, this resemblance points to a common chemical behaviour of quinuclidine and piperidine antagonists towards their environment. It supports the common binding site[33] that is proposed for CP-96,345 (**1**) and CP-99,994 (**4**).

In the crystal structures of the quinuclidine antagonists **1** and **2** and piperidine antagonists **4**-**6**, the 2-methoxy groups are not involved in any specific intermolecular interaction. Yet, internal hydrogen bonds between the 2-methoxy oxygen and the benzylamino nitrogen N2 are present in structures **1a**, **1b** and **4**. It is believed that such an internal hydrogen bond is not required for binding to the NK_1 receptor. As an alternative hypothesis to explain previously published results,[34] we propose that N2 donates a hydrogen bond to, and the 2-methoxy oxygen simultaneously accepts a hydrogen bond from the amide side chain of Gln-165 in the human NK_1 receptor. The benzylether oxygen in e.g. L-709,210 (**3**) is, on the other hand, believed to accept a hydrogen bond from Gln-165. The trifluoromethyl substituents in the benzylether group are thereby considered to be required for optimizing interactions with the human NK_1 receptor.

Amino-acid residue Gln-165 in the human NK_1 receptor is probably involved as a hydrogen bond acceptor and donor for N2 and O2, respectively, in binding the quinclidine antagonists (Fig. 13).[34] This probably also holds for the piperidine antagonists, given the common pharmacophore[16] and the similar structure-activity relationships for both classes of antagonists. Three possible sites for interactions with N2 have been located, determined with the aid of conformationally constrained analogues **5** and **6**. One position (Fig. 11) can be excluded (position A) because high affinity binding to the receptor is not lost when this position is not available for the receptor, as in antagonist **6**. Another position is shielded from the receptor by the x1 phenyl ring of the antagonists and is therefore an unlikely candidate interaction site. **The only likely position available to groups interacting with N2 corresponds to chloride position D of the piperidines (Fig. 11).** This chloride position can then probably be considered as the approximate position of the hydrogen bond donor/acceptor group in the human NK_1 receptor: the amide side chain of Gln-165. This site is geometrically described by standard hydrogen bond criteria and the dihedral angle τ(C2-

C3-N2-D), which should be in the range 130 to 164˚ (Table 6). In this argument we consider the crystal structure conformations of CP-99,994 (**4**) and **5** also energetically accessible when N2 is not protonated. Consequently, the *crystal* structure conformation of **5** is likely to be close to its bioactive conformation, given its relative rigidity in the molecular dynamics simulations and the inability of hydrogen bond formation between N2 and D in the *solution* structure of **5** (Fig.11).

A histidine residue in the human NK_1 receptor, His-265, has been reported to be involved in binding the phenyl ring of the benzylamino and benzylether group of quinuclidine antagonists (Fig. 13).[35] The crystal structures of the piperidine antagonists provide aromatic-aromatic interaction geometries for this phenyl ring. They are, however, related to the specific benzylamino conformation encountered. A non-parallel intramolecular geometry of the two phenyl rings resulted in more aromatic-aromatic interactions for the benzylamino phenyl rings in the crystal structures of CP-210,053 (**5**) and (**6**) versus CP-99,994 (**4**) (Fig. 10). This might suggest that a non-parallel orientation of phenyl rings x1 and x2 in both quinuclidine and piperidine antagonists has more possibilities for interactions in the human NK_1 receptor, which would create a favourable situation when aromatic-aromatic interactions, e.g. with His-265, are of importance. In addition, the trifluoromethyl groups of the 3,5-disubstituted phenyl rings of benzylether quinuclidine **3** and L-tryptophan benzyl ester analogue **8** seem to mediate the parallel and perpendicular stacking geometries of these phenyl rings in the two structures, respectively. This might be part of the role of the trifluoromethyl substituents in the interaction of the 3,5-disubstituted phenyl rings with His-265.[8,35]

From the class of the L-tryptophan benzyl ester antagonists a potent 3,5-bis(trifluoromethyl) analogue[36] of compound **7**, *viz.* L-732,138 was reported to bind at more or less the same binding site on the human NK_1 receptor as the benzylether quinuclidine antagonist **3**, despite their structural differences.[8] This similarity in binding is not reflected in the few intermolecular interactions observed in the crystal structures of compounds **7** and **8**, with respect to compound **3**. In structures **7** and **8** the limited functionality available in the crystal environment compared to that in a protein might be due to this. To a lesser extent this is of course also true for **3**. Additionally, it might be argued that the crystal structure conformation of **7** is, by reference to its constrained analogue **8**, probably not the bioactive one. Alternatively, the indole ring of **7** can be considered to occupy the x1' position of the

quinuclidine antagonists **1**-**3** (*vide infra*) rather than the position of phenyl ring x1, when bound to the NK_1 receptor.

Compound **9** can topologically be regarded as a hybrid of the L-tryptophan benzyl ester antagonists **7** and **8** (especially the carbonyl groups) and the benzylether quinuclidine **3** (3,5-disubstituted benzyl group and 'locked-in' benzhydryl group). Also, a comparison of the intersection volumes of antagonist **9** and CP-99,994 (**4**) in two modelled conformations has suggested that these compounds are able to occupy similar regions in space.[17] In that superposition the carbonyl oxygen O2 of **9** was located at the N2 position of CP-99,994 (**4**). However, in the crystal structure of **9** no interactions substantiating any similarity with CP-99,994 (**4**) were observed.

In analyzing intermolecular interactions in individual crystal structures, we have to be aware that they can be heavily biased due to the limited choice of functional groups in the crystal. Interaction geometries may be found only because no better alternative exists. Only interaction geometries that are in agreement with results from crystal structure statistics or high quality calculations can therefore be validly considered. Additionally, for all structures the likely difference in crystal structure conformation of flexible parts of a molecule to its bioactive conformation has to be taken into account. To understand the mechanism of recognition and binding of ligands to their receptor whenever an experimentally determined structure of a protein-ligand complex is absent, it is not sufficient to study interactions between ligands and their environment by molecular biological techniques or by analyzing intermolecular interactions in the crystal structures of the ligands. In addition to these experimental results, theoretical conformational analysis of the antagonists in relation to calculated properties indicative of interactions with the receptor environment, and the construction of binding site models is highly desirable for a more thorough understanding.

EXPERIMENTAL

Evaluation of Interactions

Hydrogen bonds were identified as such when the (potential) donor-hydrogen-acceptor angle (<D-H··A) was larger than 100.0˚, the hydrogen-acceptor distance was below the sum of their van der Waals radii minus 0.12 Å and the donor-acceptor distance was lower

than the sum of their van der Waals radii augmented by 0.5 Å.[18] Additionally, when these criteria are applied to locate C-H··X interactions (X=N, O, Cl or I usually as anions) the donor C-H group should be adjacent to a positively charged nitrogen and the hydrogen-acceptor distance should be below (or equal to) the sum of the van der Waals radii minus 0.2 Å.[20]

N^+-aromatic interaction geometries were evaluated according to the parameters defined by Verdonk et al.[29] for all aromatic rings with at least one aromatic ring carbon atom within 5 Å of any carbon atom bonded to a positively charged nitrogen.

Aromatic-aromatic interactions were classified as such when the centroids of two aromatic rings were not more than 6 Å apart, their interplane angle deviated from being perfectly parallel or perpendicular by no more than 30˚ and the smallest of the two angles between the normal of one of the rings with the vector between the centroids of the two rings was maximally 30˚.[18,37]

Tables 3, 4 and 5 have been constructed in such a way that reconstruction of the positions of the hydrogen bond acceptor groups and aromatic rings relative to the quinuclidine and piperidine rings of the antagonists is possible. In addition, files of the crystal structure conformations with all surrounding groups within 5 Å of the central ligand are available from the authors upon request.

Model Building

Molecular dynamics simulations on initial crude models of the NMR and N2 inverted conformation (configuration) of compounds **5** and **6**, respectively, were performed to obtain energetically reasonable conformations. The geometry of the initial models was first optimized to a final energy gradient 0.001 kcal·mol^{-1}·Å^{-1}, using conjugate gradients. Then a 25 ps molecular dynamics simulation was preformed at 298 K, following 5 ps equilibration. Calculations were performed with the CVFF force field in DISCOVER 2.96 (INSIGHTII 2.3.5)[38] using a dielectric constant of 1. The simulations were performed on the neutral species. The final conformations were then subjected to semi-empirical geometry optimizations performed using the AM1 Hamiltonian[24] in the MOPAC93 program.[25] Explicit definition of the geometry optimization via the eigenvector following routine, controlled by the PRECISE keyword (i.e. EF PRECISE SCFCRT=1·D-6 GNORM=0.01) was employed. MOPAC calculations were prepared and finally analyzed using SYBYL 6.2.[13]

The initial N2 inverted model of **6** was directly subjected to semi-empirical geometry optimization; molecular dynamics simulations starting with the N2 inverted model of **6** returned the crystal structure configuration every time.

ACKNOWLEDGEMENT

We thank Dr Harry R. Howard for the crystal data on CP-211,754 and Dr Richard T. Lewis and Dr Richard G. Ball for the crystal data of L-732,244 and L-709,210. Access to the PLATON program and helpful discussions on it by Dr Ton Spek are gratefully acknowledged.

REFERENCES

1. Pascard, C., *Acta Cryst.* **1995**, *D51*, 407.
2. Glusker, J. P., *Acta Cryst.* **1995**, *D51*, 418.
3. Gether, U., Lowe, J. A., III, Schwartz, T. W., *Biochem. Soc. Trans.* **1995**, *23*, 96.
4. Regoli, D., Boudon, A., Fauchere, J. L., *Pharmacol. Rev.* **1994**, *46*, 551.
5. Desai, M. C., *Exp. Opin. Ther. Patents* **1994**, *4*, 315.
6. Ward, P., Armour, D. R., Bays, D. E., Evans, B., Giblin, G. M. P., Heron, N., Hubbard, T., Liang, K., Middlemiss D., Mordaunt, J., Naylor, A., Pegg, N. A., Vinader, M. V., Watson, S. P., Bountra, C., Evans, D. C., *J. Med. Chem*. **1995**, *38*, 4985.
7. Desai, M. C., Lefkowitz, S. L., Bryce, D. K., McLean S., *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1865.
8. Cascieri, M. A., Macleod, A. M., Underwood, D., Shiao, L. L., Ber, E., Sadowski, S., Yu, H., Merchant, K. J., Swain, C. J., Strader, C. D., Fong, T. M., *J. Biol. Chem.* **1994**, *269*, 6587.
9. Allen, F. H., Davies, J. E., Galloy, J. J., Johnson, O., Kennard, O., Macrae, C. F., Mitchell, E. M., Mitchell, G. F., Smith, J. M., Watson, D. G., *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 187.
10. Howard, H. R., Shenk, K. D., Coffman, K. C., Bryce, D. K., Crawford, R. T., McLean, S. A., *Bioorg. Med. Chem. Lett.* **1995**, *5*, 111.
11. Lewis, R. T., MacLeod, A. M., Merchant, K. J., Kelleher, F., Sanderson, I., Herbert, R. H., Cascieri, M. A., Sadowski, S, Ball, R. G., Hoogsteen, K., *J. Med. Chem.* **1995**, *38*, 923.
12. Swain, C. J., Seward, E. M., Cascieri, M. A., Fong, T. M., Herbert, R., MacIntyre, D. E., Merchant, K. J., Owen, S. N., Owens, A. P., Sabin, V., Teall, M., VanNiel, M. B., Williams, B. J., Sadowski, S., Strader, C. D., Ball, R. G., Baker, R., *J. Med. Chem*. **1995**, *38*, 4793.
13. Tripos Associates, 1699 S. Hanley Road, St. Louis, Missouri 63144-2913 USA, 1995.
14. Lowe, J. A., III, Drozda, S. E., Snider, R. M., Longo, K. P., Zorn, S. H., Morrone, J., Jackson, E. R., McLean, S., Bryce, D. K., Bordner, J., Nagahisa, A., Kanai, Y., Suga, O., Tsuchiya, M., *J. Med. Chem.* **1992**, *35*, 2591.

15. Snider, R. M., Constatine, J. W., Lowe, J. A., III, Longo, K. P., Lebel, W. S., Woody, H. A., Drozda, S. E., Desai, M. C., Vinick, F. J., Spencer, R. W., Hess, H. J., *Science* **1991**, *251*, 435.
16. Desai, M. C., Lefkowitz, S. L., Thadeio, P. F., Longo, K. P., Snider, R. M., *J. Med. Chem.* **1992**, *35*, 4911.
17. Natsugari, H., Ikeura, Y., Kiyota, Y., Ishichi, Y., Ishimaru, T., Saga, O., Shirafuji, H., Tanaka, T., Kamo, I., Doi, T., Otsuka, M., *J. Med. Chem.* **1995**, *38*, 3106.
18. Spek, A. L., *Acta Cryst.* **1990**, *A46* , C34.
19. Burley, S. K., Petsko, G. A., *Adv. Prot. Chem.* **1988**, *39*, 125.
20. Taylor, R., Kennard, O., *J. Am. Chem. Soc*. **1982**, *104*, 5063.
21. Allen, F. H., *Acta Cryst.* **1986**, *B42*, 521.
22. Bernstein, J., Davis, R. E., Shimoni, L., Chang, N.L., *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1555.
23. Desai, M. C., Vincent, L. A., Rizzi, J. P., *J. Med. Chem.* **1994**, *37*, 4263.
24. Dewar, M. J. S., Zoebisch, E. G., Healy, E. F., Stewart, J. P., *J. Am. Chem. Soc.* **1985**, *107*, 3902.
25. Stewart, J. P., *J. Comput.-Aided Mol. Design* **1990**, *4*, 1.
26. Macleod, A. M., Merchant, K. J., Brookfield, F., Kelleher, F., Stevenson, G., Owens, A. P., Swain, C. J., Cascieri, M. A., Sadowski, S., Ber, E., Strader, C. D., MacIntyre, D. E., Metzger, J. M., Ball, R. G., Baker, R., *J. Med. Chem.* **1994**, *37*, 1269.
27. Lowe, J. A., III, Drozda, S. E., McLean, S., Crawford, R. T., Bryce, D. K., Bordner, J., *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1153.
28. Dougherty, D. A., *Science* **1996**, *271*, 163.
29. Verdonk, M. L., Boks, G. J., Kooijman, H., Kanters, J., Kroon, J., *J. Comput.-Aided Mol. Design* **1993**, *7*, 173.
30. Steiner, Th., Starikov, E. B., Amado, A. M. and Texeira-Dias, J. J. C., *J. Chem. Soc., Perkin Trans.* **1995**, *2*, 1321.
31. Gether, U., Nilsson, L., Lowe, J. A., III, Schwartz, T. W., *J. Biol. Chem.* **1994**, *269*, 23959.
32. Fong, T. M., Cascieri, M. A., Yu, H., Bansal, A., Swain, C., Strader, C. D., *Nature* **1993**, *362*, 350.
33. Pradier, L., Habert-Ortoli, E., Emile, L., Le Guern, J., Loquet, I., Bock, M. D. , Clot, J., Mercken, L., Fardin, V., Garret, C., Mayaux, J. F., *Mol. Pharmacol.* **1995**, *47*, 314.
34. Fong, T. M., Yu, H., Cascieri, M. A., Underwood, D. , Swain, C. J., Strader, C. D., *J. Biol. Chem.* **1994**, *269*, 14957.
35. Fong, T. M., Yu, H., Cascieri, M. A., Underwood, D. , Swain, C. J., Strader, C. D., *J. Biol. Chem.* **1994**, *269*, 2728.
36. Macleod, A. M., Merchant, K. J., Cascieri, M. A., Sadowski, S., Ber, E., Swain, C. J., Baker, R., *J. Med. Chem.* **1993**, *36*, 2044.
37. Hunter, C. A., Sanders, J. K. M., *J. Am. Chem. Soc.* **1990**, *112*, 5525.
38. Biosym Technologies, 9685 Scranton Road, San Diego, CA 92121-2777 USA, 1995.

Chapter 3

Electrostatic Complementarity between NK_1 Antagonists and Their Crystal Environment: A Model for Ligand-Receptor Interactions

ABSTRACT

Extended electron distributions (XEDs) have been used to study the electrostatic complementarity in the crystal structures of non-peptide NK_1 antagonists and the possible consequences thereof for binding to the human NK_1 receptor. The common anion/hydrogen bond acceptor sites in the crystal structures of quinuclidine NK_1 antagonists were found to be in good correspondence with positive extrema positions on the XED molecular electrostatic-potential (MEP) in different ion-pair complexes of the *N*-methyl analogue of the prototypic NK_1 antagonist CP-96,345. Remarkably, the influence of different counter-ion positions on the resulting MEP extrema positions was only minor. In addition, N^+-aromatic and aromatic-aromatic stacking interaction geometries, as have been encountered in the antagonist crystal structures, were in agreement with favourable positions for a benzene molecule after multiple dockings to various ion-pair complexes of the CP-96,345 analogue. Given this correspondence, it is concluded that the use of extended electron distributions in studying ligand-receptor interactions is a promising approach.

INTRODUCTION

In Chapter 2 we analyzed the intermolecular interactions in crystal structures of non-peptide NK_1 antagonists. The interaction geometries encountered have been found to correspond to reported ligand-receptor interactions and to the structure-activity relationships of these compounds. Therefore, the interactions identified might reflect possible interaction geometries with the human NK_1 receptor. Only interaction geometries could be identified that corresponded to established interaction criteria, either inferred from crystal structure statistics or from high quality quantum chemical calculations; other interactions might only be present in the crystal structures because no better alternative existed. In this respect, it is important to realize that the variety of functional groups in small-molecule crystal structures is severely limited compared to the functionality present in protein binding sites. The interactions encountered in individual small-molecule crystal structures will therefore be far from representative for all possible interactions of a ligand. In addition, the interactions found are associated with the ligand conformation observed, which is usually not its bioactive conformation except for extremely rigid molecules. In this Chapter the usefulness of a set of mutually complementary empirical methods is assessed that can be helpful in addressing these problems.

Vinter developed empirical methods which generalize the study of possible interactions of ligands, taking conformational flexibility into account.[1a-c] The methods are based on fields generated by approaching molecules (*viz*. electrostatic, dispersive and inductive fields), that determine the interactions between a molecule and its environment. Essential to these methods is a molecular mechanics force field that better accounts for the electron distribution in a molecule. Following the concept of using multi-centred charges Vinter developed the so-called extended electron distributions (XEDs) for the description of the charge distribution of atoms (see Methods). With multi-centred point charges the partial charges are assigned to positions around the nucleus that roughly mimic the natural atomic orbitals.[2] In addition, XEDs have been parametrized to reproduce experimentally or quantum-chemically determined geometries of binary complexes.[1c] Thereby, polarization is implicitly taken into account. XEDs have been found to be transferable for the calculation of other intermolecular interactions,[1a,c] and to reproduce the molecular electrostatic-potential (MEP) extrema generated by high-level quantum-chemical distributed multipole analyses.[1b,3] It has been shown that molecules that bind at similar binding sites, can relatively successfully be superimposed by means of optimizing the similarity of electrostatic properties (including electrostatic-potential extrema), but success has been shown to be

highly dependent on the quality of the electrostatic model.[4-7] In this respect, multi-centred point charges, instead of the commonly used atom-centred charges, have resulted in an improved description of the MEP[2] and aromatic-aromatic interaction geometries.[8] These methods are, in principle, not restricted to specific types of interactions or to one conformation. In the end, it might be possible to create a molecular 'fingerprint' of a ligand that describes the molecular recognition of that ligand and a given macromolecular receptor. It has already been shown that regioselectivity of CCK-B antagonists can be tackled by these methods.[9] Also, experimentally determined geometries for the stacking of e.g. porphyrin rings have been correctly predicted,[1c] which had not been possible using conventional molecular mechanics approaches.[8] These methods were assessed with respect to their ability to reproduce recurring hydrogen bond acceptor/anion positions as well as aromatic interaction sites for quinuclidine NK_1 antagonists in their crystal environment.[10] Especially the influence of surrounding anions on the intermolecular interactions was evaluated. The work reported may be helpful in the development of a protocol for the semi-quantitative identification and interpretation of characteristic features of molecular fields in terms of possible interaction geometries.

METHODS

XED Electrostatic and Surface Potential Extrema

The electron distribution of the molecules was described using extended electron distributions (XEDs) on all heteroatoms and non-sp3 hybridized carbon atoms.[1c] In the application of XEDs, the total electron density allocated to a single atom is distributed to positions outside the nucleus, leaving the atomic charge unchanged. In this way electronic anisotropy is partially taken into account. Atom-centred partial charges were allocated to all other atoms.

In Vinter's perception, the preferential interactions of a molecule can be approximated by important features of the molecular fields for the electrostatic, van der Waals and induction components of intermolecular interactions.[1b] These important features include extrema of the molecular electrostatic potential (MEP) close to the molecular van der Waals surface as well as sites were the polarization of an approaching probe molecule would be (locally) maximal.[1b] The extrema are calculated by separately minimizing the interaction energy of a unitary positively or negatively charged probe, as well as a polarizable probe[1b] starting from various points at (close to) the molecular surface of the central ligand. The

electrostatic part of the interaction of the unitary charged probes with the XED electron density is described by a Coulomb potential summed over all partial charges (atoms and XEDs). The induction part of the interaction energy of the polarizable probe (in kcal/mol) is given by Eq.1,

$$E_I = 0.5 * \alpha_p * \sum q_i / Dr_{ip}^4 \qquad (1)$$

where α_p represents the static polarizability of the probe. Here, the polarizability of water α_p=3 cm^3 was used.[11] The summation is over all partial charges q_i (atoms and XEDs) in the molecule, at a distance r_{ip} to the probe.

The dielectric constant (D) equals 3 in all electrostatic interaction terms. Van der Waals interactions are incorporated by adding a Morse-potential to the potential terms describing the electrostatic interaction of the probes, employing an oxygen van der Waals radius for the probes. Minima (-ve) and maxima (+ve) on the XED electrostatic potential are positions where the interaction energies for the positively and negatively charged probe are in a minimum, respectively. For the polarizable probe only sites with a total interaction energy below -10.0 kcal/mol were taken into account; sites with interaction energies above that value were found to be highly dependent on the orientation of the molecule in coordinate space. Note that, in contrast to the method as originally described in ref. 1b, no further scaling of the interaction energy of the polarizable probe has been performed.

XED Dockings

Dockings of a rigid probe molecule (the mobile species) to a fixed target molecule were performed in the program XEDOCK by placing the mobile probe on evenly spaced points on a sphere with the target molecule at its centre.[1c] After orienting the probe molecule at its initial position to determine the most favourable starting orientation, a six-dimensional simplex minimization (rotation and translation) of the non-bonded interaction energy between probe and target is performed. Typically 259 starting points for the probe on the sphere were employed in each docking run. In this case the mobile probe was a benzene molecule which was, as well as the target molecule, allocated with XEDs. The radius of the sphere is dependent on the target molecule and varied between 16.8 and 18.6 Å in our calculations. The dielectric constant employed in the XED dockings was either D=1 or D=3.

In employing crystal structure conformations both in the calculation of potential extrema as well as in the dockings, hydrogen atom positions were recalculated using SYBYL[12] to normalize the crystallographically determined bond lengths involving hydrogen atoms.[13] XEDs and XEDOCK have been incorporated in the COSMIC molecular modelling package.[14,15]

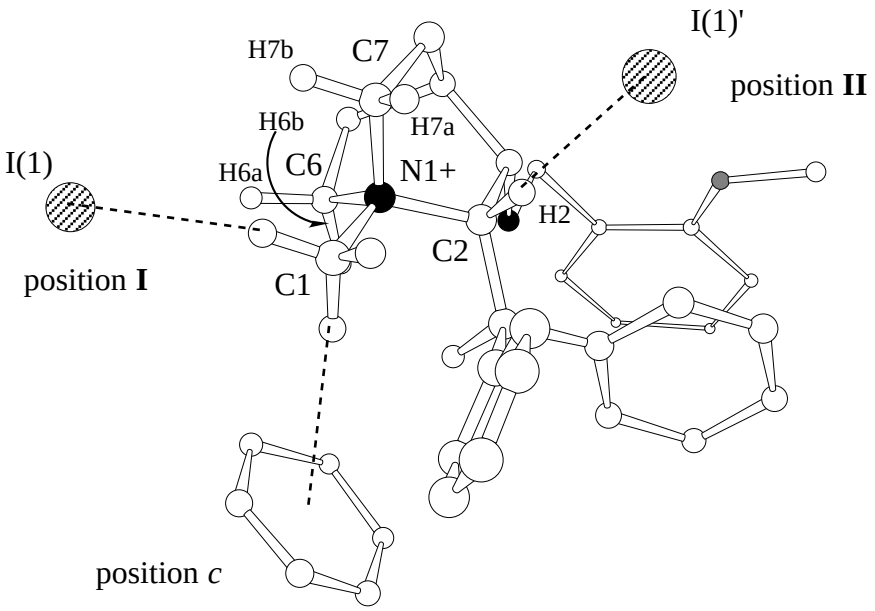


FIGURE 1. Intermolecular interaction geometries in the crystal structure of the *N*-methyl analogue of CP-96,345 (*viz.* LEWCUL).[10,16] See also Chapter 2 of this thesis.

CRYSTAL STRUCTURE OF LEWCUL

The intermolecular interactions in the crystal structure of the *N*-methyl analogue of the non-peptide NK_1 antagonist CP-96,345, henceforward identified by its CSD refcode LEWCUL, are depicted in Fig. 1.[10] LEWCUL carries a permanent positive charge and has been crystallized as the monoiodide salt.[16] The iodide positions near the quinuclidine ring are common anion interaction sites for the positively charged quinuclidine antagonists and have been denoted as positions I and II.[10] Iodide anion positions I and II for LEWCUL can be described in terms of C-H··I^- (iodide) interactions arising from C1-H and C2-H2, respectively (Fig. 1). The phenyl ring position *c* (Fig. 1) is stabilized by both N^+-aromatic interactions and C-H··π interactions[17] donated by the benzhydryl (diphenylmethyl) group.[10] The benzhydryl group coordinates a methoxymethyl group, probably by hydrophobic

interactions (data not shown). The crystal structure conformation of LEWCUL with corrected hydrogen positions (see Methods) was used in subsequent calculations.

RESULTS

XED Electrostatic-potential Extrema

XED electrostatic-potential extrema surrounding LEWCUL in its crystal structure conformation were evaluated for different ligand-anion complexes, in which model anions occupy the respective crystallographic iodide positions (Fig. 1). The model anions were given a unitary negative charge and an oxygen van der Waals radius. Correspondingly, the net formal charge of the ion-pairing complexes ranges from +1 (**1**, no anions) via 0 (one anion either in position I (**2**) or II (**3**)), to -1 (**4**, two anions), as depicted in Fig. 2.

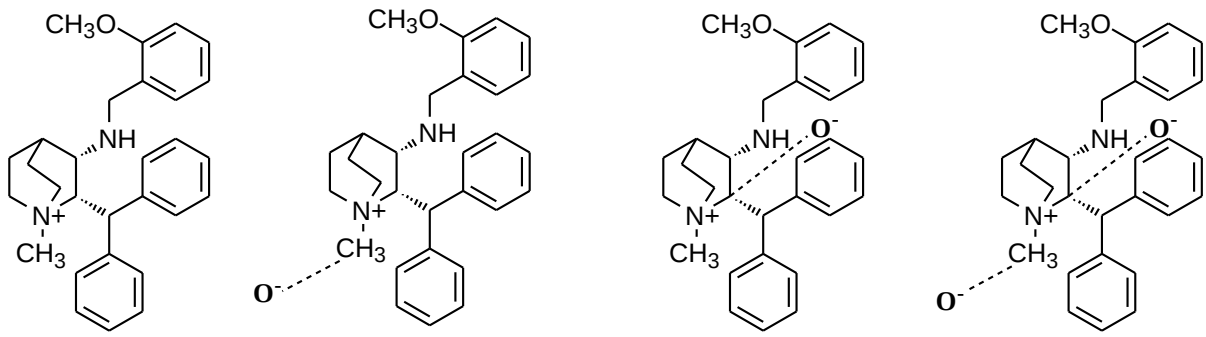


FIGURE 2. Schematic representation of the ligand-anion complexes included in the calculations.

The extrema positions in the positively charged species (**1**) are depicted in Fig. 3. Positive extrema (+ve) were present around the quinuclidine ring, at the edge of the benzylamino phenyl ring and near the methoxy methyl group at roughly equal distances to the two neighbouring phenyl rings of LEWCUL. Obviously, negative extrema (-ve) are not present for this species. Two favourable positions for the polarizable probe were located, one in the

TABLE 1. Corresponding anion and positive XED electrostatic-potential extrema positions (+ve) for the crystal structure conformation of LEWCUL in various ion-pairing complexes

Position I						
Ion-pairing complex[a]	Donor	Acceptor/ extremum[b]	d(N1···+ve) (Å)	<(C2-N1···+ve) (˚)	τ(C2-N1-C1·+ve) (˚)	τ(N1-C2-C3·+ve) (˚)
	C1-H	**I(1)**	**4.652**	**165.5**	**-172.7**	**-6.8**
1 (+1)	C1-H	+ve(131)	3.32	174	178	2
2 (I,0)	-	-	-	-	-	-
3 (II,0)	C1-H	+ve(137)	3.40	174	-179	0
Position II						
Ion-pairing complex[a]	Donor	Acceptor/ extremum	d(N1···+ve) (Å)	<(N1-D···+ve) (˚)	τ(N1-D-H·+ve) (˚)	τ(N1-C2-C3·+ve) (˚)
	C2-H2	**I(1)'**	**4.845**	**111.0**	**-109.3**	**113.2**
1(+1)	C2-H2	+ve(132)	3.60	94	-22	100
	C7-H7a	+ve(132)	3.60	97	74	100
2 (I,0)	C2-H2	+ve(138)	3.96	112	-115	114
	C7-H7a	+ve(138)	3.96	92	75	114
3 (II,0)	-	-	-	-	-	-

Crystallographic anion positions and atom labelling as given in Fig. 1. D denotes a donor. a) In parentheses in the first column, the net formal charge and the position that is actually occupied by an anion, respectively. b) The positive extrema are indicated by identifiers from the molecular data files. Corresponding crystallographic values for I(1) and I(1') as acceptor are given in bold (see also Table 3, Chapter 2).

pocket formed by the phenyl rings of the benzhydryl group, the other close to the lone pair of N2 (Fig. 3). The three positive extrema around the quinuclidine ring correspond to the positions of the anions in position I and II and the phenyl ring in position *c* (Tables 1 and 2). These positive extrema were also present (but less pronounced) in the ion-pair complexes, except of course the site that was actually occupied by an anion in an individual complex (Tables 1 and 2). In the complex in which both positions I and II were occupied by anions, as can be expected, no positive extrema were found. Values for the electrostatic potential at the positive extrema in the various complexes are given in Table 3. Potential values in the ion pairs were approximately half the values of the positive extrema in the net positively charged species.

TABLE 2. Positive XED electrostatic-potential extrema (+ve) and the corresponding aromatic ring position in the crystal structure of LEWCUL

Crystall. position	C(N+)	Ring position	d(N1..x) (Å)	d(C(N+)..x) (Å)	<(C(N+)-N1..x)[a] (˚)
	C1	*c*	5.292	5.032	71.9
	C6	*c*	5.292	4.749	61.3
Ion-pairing complex	C(N+)	+ve Extremum[b]	d(N1..+ve) (Å)	d(C(N+)..+ve) (Å)	<(C(N+)-N1..+ve) (˚)
1 (+1)	C1	129	3.45	3.18	67
	C6	129	3.45	3.03	61
2 (I,0)	C1	139	4.16	4.25	83
	C6	139	4.16	3.69	62
3 (II,0)	C1	138	3.52	3.24	67
	C6	138	3.52	3.07	61

a) This angle is the angle θ that is used for describing the N^+-aromatic interaction geometries (see Chapter 2, Table 4). x denotes the ring centroid. b) The numbers are identifiers for the various extrema

The favourable positions for a polarizable probe near the positively charged species reappeared at similar positions in all three ion-pairing complexes. As an example, in the positively charged species a site was located at 2.9 Å from N2 (Fig. 3) and another at 3.5 Å from both centroids in the pocket formed by the phenyl rings of the benzhydryl group.

In the two neutral and the net negatively charged ion-pairing complexes, the most pronounced negative extrema were clearly paired with the anions in the complexes and mimic a second coordination shell (results not shown). Less pronounced, but recurring negative extrema were found near the π-electron clouds of the exposed face of the benzylamino phenyl ring and at the centre of the pocket formed by the phenyl rings of the benzhydryl group (called the benzhydryl pocket from hereon). In the net negatively charged complex also a negative extremum was present at the side of the benzylamino phenyl ring that faces the benzhydryl group. Typically, the negative extrema in complexes 2, 3 and 4 were found directly above or below the ring centroid at distances ranging from 3.1 to 3.4 Å. Potential values differ in the various complexes, but in all cases the negative extremum value in the benzhydryl pocket was roughly twice the value of the extremum near the benzylamino phenyl ring in the same complex (Table 4).

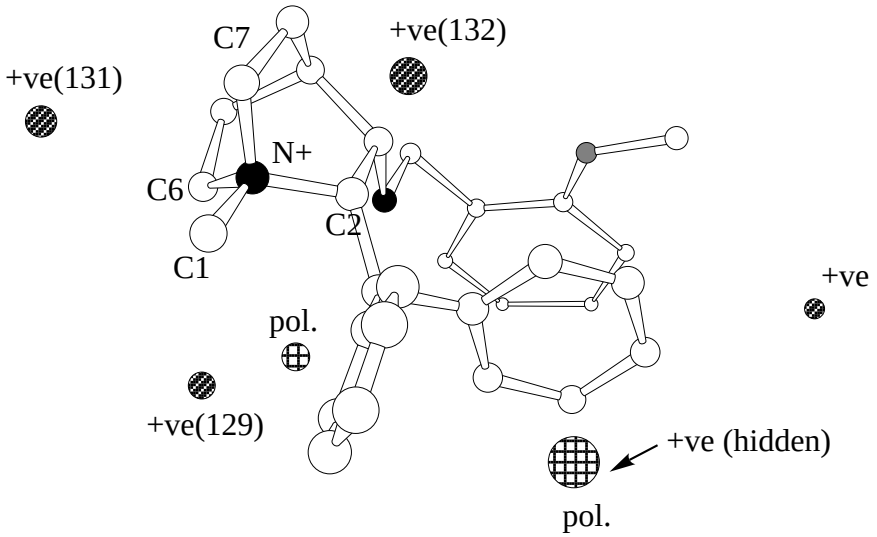


FIGURE 3. XED electrostatic-potential extrema exhibited by the net positively charged species of LEWCUL. The radii of the spheres are proportional to the (absolute) value of the potential at that point. 'pol.' designates a polarizable probe

TABLE 3. XED electrostatic-potential values (in kcal/mol) at recurring extrema positions around the quinuclidine ring of LEWCUL for different ion-pairing complexes

Ion-pairing complex	Corresponding crystallographic positions		
	I	**II**	*c*
1 (+1)	30.5	33.0	30.8
2 (I,0)	anion	18.4	11.3
3 (II,0)	15.6	anion	16.5

Ion-pairing complexes as described in Tables 1 and 2. The extrema are identified by the corresponding crystallographic positions.

The positions occupied with anions surrounding the quinuclidine rings in the antagonist crystal structures have been ascribed C-H··X (X=O, N, chloride or iodide) interactions.[10,18] To further investigate their relation with positive extrema positions, these positions were evaluated in terms of parameters that are used in describing C-H··X interactions (Table 5). The parameter *d* in Table 5 gives the shortening of the interatomic distance with respect to the distance at van der Waals contact.

TABLE 4. Negative XED electrostatic-potential extrema (-ve) at recurring positions in ion-pairing complexes of LEWCUL

Ion-pairing complex	Associated group	d(x··-ve) (Å)	Potential value (kcal/mol)
2 (I,0)	benzylamino phenyl	3.1	-2.7
	benzhydryl	3.2	-5.4
3 (II,0)	benzylamino phenyl	3.1	-5.3
	benzhydryl	3.4	-12.7
4 (I,II,-1)	benzylamino phenyl	3.1, 3.3	-10.9, -14.3
	benzhydryl	nd	-23.9

Ion-pairing complexes as described in footnote a in Table 1. nd is not determined.

XED Dockings

Multiple dockings (see Methods) of a benzene molecule carrying XEDs to the net neutral LEWCUL ion-pairs (also carrying XEDs) were performed. Dockings were performed using a dielectric constant of D=1 or D=3. The *in vacuo* value (D=1) has been employed in the parametrization of the XEDs,[1c] whereas D=3 is consistent with the calculation of the XED electrostatic-potential extrema. Resulting configurations with interaction energies within 3.0 kcal/mol from the global interaction energy minimum in each run are shown in Fig. 4. The resulting docked configurations over the *entire* energy ranges did not differ considerably as far as the region around the quinuclidine rings was concerned (data not shown); the main differences were related to the relative interaction energies of the clusters of docked benzene rings, as can be seen in Fig. 4. In Figs 4 (a), (b) and (d) the positions of the benzene rings with the lowest (negative) interaction energy (global minimum, e.g. cluster 1 in Fig. 4 (b)) more or less correspond to position *c* in the crystal structure of LEWCUL. The benzene ring clusters in Fig. 4 (b), that resulted from the dockings to the ion-pair with an anion in position I and a dielectric constant D=3, were evaluated in terms of N^+-aromatic interaction geometries[20] (Table 6); cluster members with the lowest interaction energy have been taken to represent each cluster.

Aromatic-aromatic interaction geometries occurring within 3.0 kcal/mol of the global interaction energy minimum, were only observed for the benzhydryl group in Fig. 4 (d). At higher interaction energies also benzene rings in the vicinity of the benzhydryl group in other

TABLE 5. Positive XED electrostatic-potential extrema positions (+ve) surrounding the quinuclidine ring of NK_1 antagonist LEWCUL for different ion-pairing complexes, evaluated as C-H··X interactions[18]

Crystall. positions	D-H··Ac	d(D··Ac) (Å)	d(D-H) (Å)	d(H··Ac) (Å)	<(D-H··Ac) (°)	*d* (Å)[b]
I	C1-H··I(1) iodide	3.902	1.10	2.82	167	-0.36
II	C2-H2··I(1)' iodide	4.071	1.10	3.01	163	-0.17

Ion-pairing complexes

No.	Donor	+ve extremum	Crystal position	d(D··+ve) (Å)	d(H··+ve) (Å)	<(D-H··+ve) (°)	*d* (Å)
1 (+1)	C1-H	131	**I**	3.15	2.22	141	-0.50
	C7-H7b	131	**I**	3.10	2.16	142	-0.56
	C6-H6a	131	**I**	3.10	2.22	135	-0.50
	C2-H2	132	**II**	3.14	2.09	159	-0.63
	C7-H7a	132	**II**	3.09	2.27	129	-0.45
	C1-H	129	*c*	3.18	2.26	139	-0.46
	C6-H6b	129	*c*	3.03	2.22	128	-0.50
2 (I,0)	C2-H2	138	**II**	3.11	2.05	160	-0.67
	C7-H7a	138	**II**	3.62	3.04	113	0.32
	C1-H	139	*c*	4.25	3.38	136	0.66
	C6-H6b	139	*c*	3.69	2.64	159	-0.08
3 (II, 0)	C1-H	137	**I**	3.21	2.27	142	-0.45
	C6-H6a	137	**I**	3.11	2.18	140	-0.54
	C7-H7b	137	**I**	3.24	2.31	141	-0.41
	C1-H'	138	*c*	3.24	2.32	140	-0.40
	C6-H6b	138	*c*	3.07	2.25	129	-0.47

Crystallographic positions and atom labelling as given in Fig. 1. D, Ac and +ve are donors, acceptors and positive extrema positions, respectively. Extrema are indicated by identifiers from the molecular data files. All C-H bond lengths are 1.10 Å b) The parameter d = d(H··Ac) - R_{vdW}(H) - R_{vdW}(Ac); R_{vdW}(H)=1.20 Å. R_{vdW}(O)=1.52 Å is taken as the probe radius with respect to the extrema, for the iodide ion R_{vdW}(I)=1.98 Å. Van der Waals radii are from Bondi et al.[19]

dockings appeared, as shown in Fig. 5. Aromatic-aromatic stacks involving the benzylamino phenyl ring were observed at energies well above 3.0 kcal/mol from the global minimum (data not shown).

An additional benzene docking run was performed on the crystal structure conformation of the 3,5-bis(trifluoromethyl)benzyl ether analogue of CP-96,345 (*viz.* L-709,210)[21,10] as a neutral ion-pair (anion in position I). The global minimum interaction energy corresponds to the configuration with a benzene ring near position *c* of the crystal structure of LEWCUL. The disubstituted benzyl group was with either face of the phenyl ring (independently) involved in face-to-face aromatic stacks with benzene rings (Fig. 6). Benzene-benzyl intercentroid distances are 3.5 Å for both benzene rings. These configurations were present within 3.0 kcal/mol from the global interaction energy minimum.

DISCUSSION

The positions of the positive electrostatic-potential extrema (+ve) in the positively charged species of LEWCUL were virtually conserved in the neutral ion-pair complexes with the exception of the position that was actually occupied by an anion (Tables 1 and 2). Only the height of the extrema. i.e. the potential values, varied (Table 3). Comparison of the positive extrema positions with the anion and aromatic ring positions in the crystal structure of LEWCUL revealed a remarkable agreement (Tables 1 and 2), largely irrespective of the nature of the ligand, either net positively charged or as an ion-pair. The agreement is most pronounced in the angular parameters that were used to describe the extrema positions. Not surprisingly, the extrema-to-N1 and extrema-to-C(N+) distances differed from the iodide-to-N1 and phenyl centroid-to-C(N+) distances, due to the difference of the probe size and probe shape with that of the iodide anion and the phenyl ring in the crystal structure. This overall agreement of positive electrostatic-potential extrema and experimentally observed (recurring) anion and aromatic ring positions can be considered to illustrate electrostatic (and steric) complementarity.

The results given in Table 5 suggest that positive extrema positions may be understood in terms of C-H··X interactions,[17,18] in which an anion or the π-electron cloud of a phenyl ring fulfils the acceptor role of X. They also suggest that N^+-aromatic interactions[22,20] for some ligands might be understood in terms of C-H··X interactions. Although the C-H groups adjacent to the positively charged nitrogen of the quinuclidine ring geometrically fulfil a donor role in C-H··X interactions in the crystal structures (see Chapter 2), the nature of the interaction may well be different from e.g. C-H··O hydrogen bond interactions.[23] High-quality *ab initio* quantum chemical calculations on C-H··O hydrogen

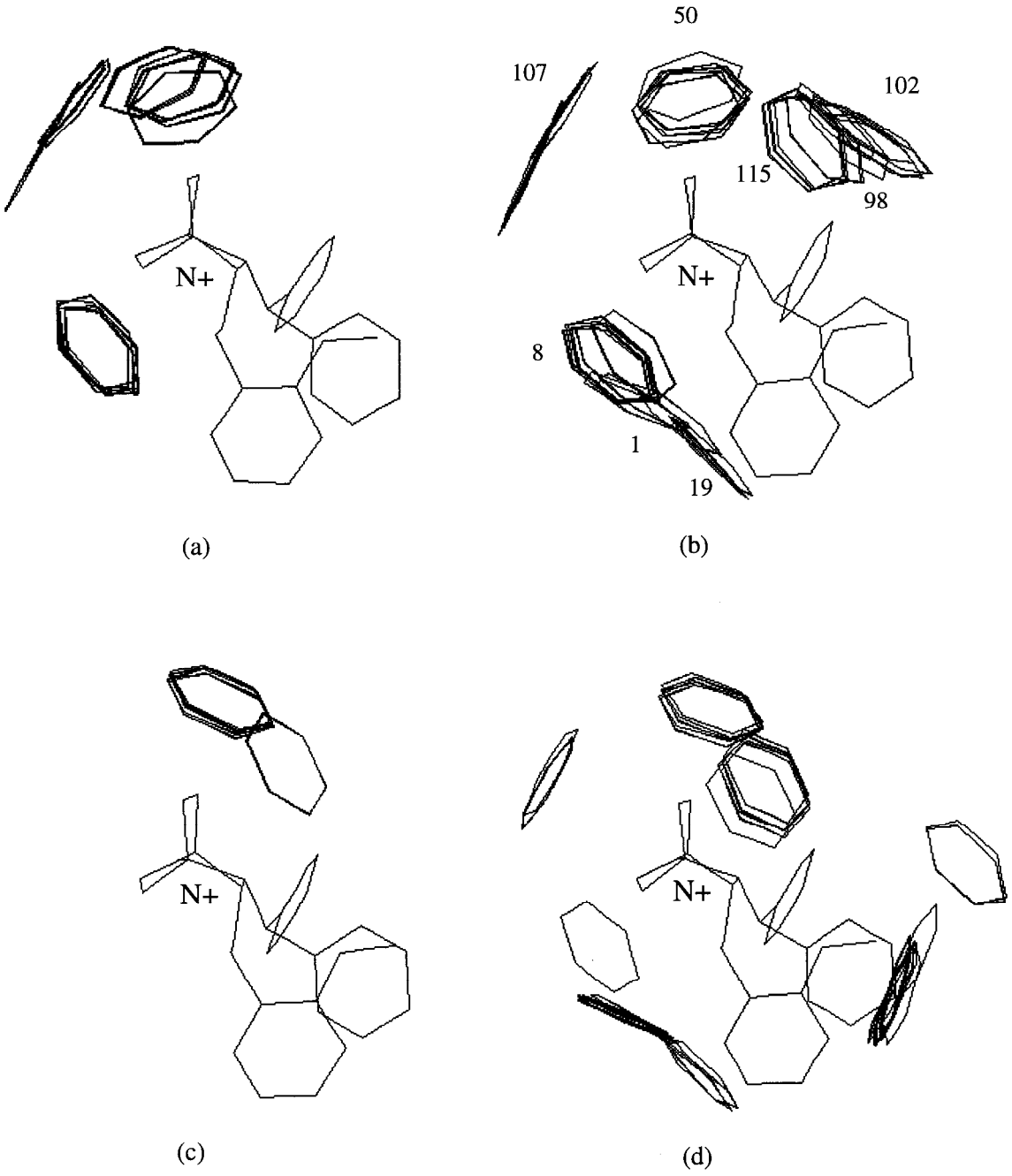


FIGURE 4. Resulting geometries after multiple dockings of a benzene molecule to the various ion-pair complexes of LEWCUL. The complexes are depicted in projection along the C1-N1+ bond (Fig. 1). Only depicted are benzene rings with interaction energies within 3.0 kcal/mol of the lowest interaction energy in that run. In the calculations an anion was present in position I with (a) D (dielectric constant) = 1.0 and (b) D=3.0, or in position II with (c) D=1.0 and (d) D=3.0.

TABLE 6. Representative interaction geometries for clusters of benzene rings after docking to ion pair **2** (I,0) of LEWCUL (D=3)

Ring position[a]	C(N+)	d(N1··x) (Å)	d(C(N+)··x) (Å)	<(C(N+)-N1··x)[b] (˚)	ϕ[b] (˚)	τ(C2-N1-C(N+)··x) (˚)
1[c]	C1	4.36	4.50	85	16	-87
	C6	4.36	3.44	45	16	66
50	C1	4.19	3.36	47	12	108
	C7	4.19	3.69	61	12	-115
98[d]	C2	5.01	4.18	50	26	-
	C7	5.01	4.62	67	26	-30
107	C6	4.53	4.53	80	21	-139
	C7	4.53	3.41	35	21	159

a) As depicted in Fig. 3. Numbering is sequential in accordance with the energy-sorted docking positions: 1 corresponds to the global minimum. x denotes the ring centroid. b) These angles correspond to the angles θ (<(C(N+)-N+..x)) and ϕ that are used for describing N+-aromatic interaction geometries.[20] ϕ is the angle from N1, *via* the ring centroid x, to the endpoint of the normal of the plane of the benzene ring directed towards N1. c) 1 represents clusters 1, 8 and 19. d) 98 represents clusters 98, 102 and 115.

bonded systems, including the binary CH_3-NH_3^+··H_20 complex, revealed a large shortening of the equilibrium C··O distance (2.91 versus 3.81 Å) when the donor carbon atom was positioned adjacent to a positively charged nitrogen atom.[24] The relatively small exchange repulsion of the methyl carbon atom despite these short distances was overcome by the electrostatic attraction between the NH_3^+ group and the (aligned) dipole of the water molecule.[24] This suggests that both the anion positions in the crystal structure of LEWCUL and other quinuclidine antagonists,[10] as well as the positive extrema positions can be rationalized in terms of the electrostatic N^+-anion/probe attraction for which the C-H groups oppose a closer approach.

The interaction geometries of the benzene rings surrounding the quinuclidine ring in the various LEWCUL ion-pairing complexes were found to be conserved to a considerable degree. In addition, the benzene rings belonging to these clusters are involved in N^+-aromatic interaction geometries[20] (Table 6). The benzene rings were always encountered in interaction geometry type 1, which involves two different C-H groups adjacent to N+ (Table 6).[20] In the crystal structures of the positively charged quinuclidine antagonists (Table 4 in

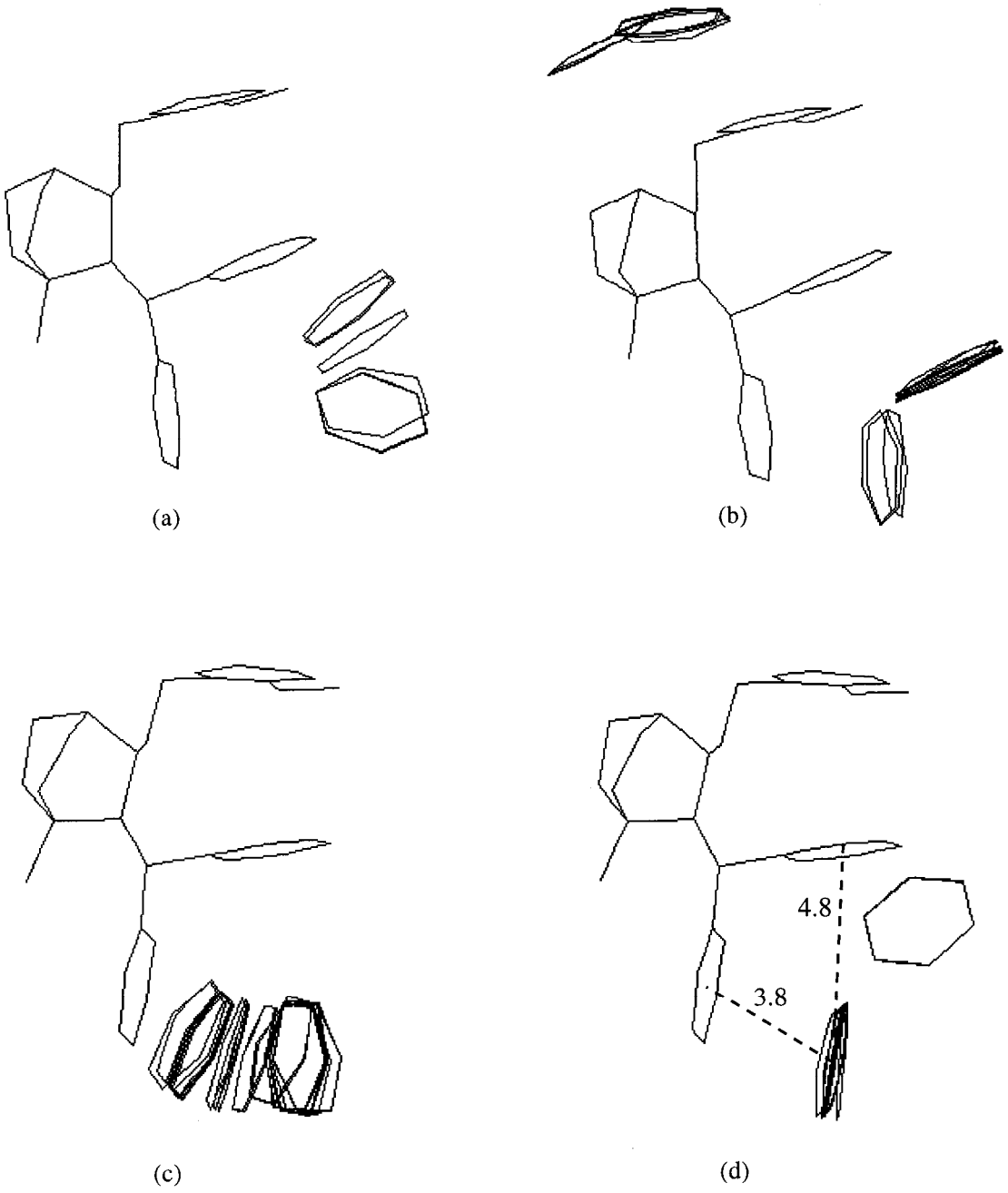


FIGURE 5. Typical interaction geometries of benzene rings with the benzhydryl group resulting from the dockings described in Fig. 3. Various interaction energy ranges[29] are shown for the various ion-pair complexes. (a) Anion in I, D=1, energy range -8.5 to -8.0 kcal/mol. (b) Anion in I, D=3, energy range -7.3 to -6.5 kcal/mol. (c) Anion in II, D=1, energy range -9.7 to -7.9 kcal/mol. (d) Anion in II, D=3, energy range -7.4 to -7.2 kcal/mol.

Chapter 2) also the geometry type involving only one C-H group (type 2 in Chapter 2 and ref. 20), *viz*. positions *a*, *b* and *d* (Fig. 6 in Chapter 2), were present. On the other hand, in the dockings benzene rings were surrounding the entire quinuclidine ring, whereas in the crystal structures only one hemisphere of the quinuclidine ring was occupied. These differences can probably be attributed to the other groups constituting the crystal structure, whereas the benzene rings were docked in isolation. By analogy with the results for the potential values at the positive extrema (Table 3), the main differences between the benzene ring clusters in the various complexes were associated with the varying relative interaction energies of similar clusters in the various complexes (Fig. 4). Yet, ring positions associated with each numbered cluster in Fig. 4 (b) reappeared in at least one other docking run within 3.0 kcal/mol from the global minimum interaction energy in that run. In addition, some, but not all, clusters (i.e. clusters 1, 98 and 107) occupied sites where positive potential extrema were also located. This indicates that anions (that occupy the positive extrema) and aromatic rings interacting with positively charged groups have important features in common. Furthermore, N^+-aromatic interaction geometries are largely conserved irrespective of the specific ion-pair or dielectric constant, which stresses the importance of these interactions in this case.

In the benzene dockings on the neutral ion pair of the benzylether analogue of CP-96,345, *viz.* L-709,210,[21] position *c* of the LEWCUL crystal structure again was the site with the lowest (negative) interaction energy. The methyl group at N1 in LEWCUL is therefore considered not to be a prerequisite for favourable N^+-aromatic interactions at position *c*. Also C-H··π interactions involving the benzhydryl group seem to play a role. It supports the hypothesis[10] that position *c* is a favourable and ligand-specific interaction site for aromatic rings at the positively charged quinuclidine ring in NK_1 antagonists, and therefore perhaps also in interactions with the human NK_1 receptor. Possible candidates are Phe-268 or Tyr-272 in that receptor, as is schematically depicted in Fig. 13 in Chapter 2.

Negative electrostatic potential extrema (-ve) appeared at typical distances of 3.3 Å from the centroid of phenyl rings, near the π-electron clouds. The extremum in the benzhydryl pocket is due to the cooperative action of the two phenyl rings. This is reflected in the negative potential value at that site, which is roughly twice the potential value at the recurring extremum near the benzylamino phenyl ring (Table 4). It is also readily understood that an anion in position II (Fig. 1) which is closer to the phenyl rings than the anion in position I, will to a larger extent affect the potential near these phenyl rings (Table 4). In the

crystal structure of LEWCUL no specific intermolecular interactions have been found for the benzhydryl group. Yet, in the benzene docking results in the neutral LEWCUL ion-pair complexes with D=3 edge-to-face and face-to-face aromatic stacks (Figs 5 (b) and (d)) similar to the ones found in the crystal structure of the dimesylate salt of CP-96,345 (Chapter 2, Fig. 8 (b)) were identified. The docked benzene rings near the benzhydryl group of the neutral ion-pairs calculated with a dielectric constant D=1, on the other hand, did not reveal stacked geometries. In contrast to the remarkable similarity in the benzene cluster positions and orientations surrounding the quinuclidine ring in the four docking runs, the interaction geometries near the benzhydryl group were found to be dependent on the dielectric constant

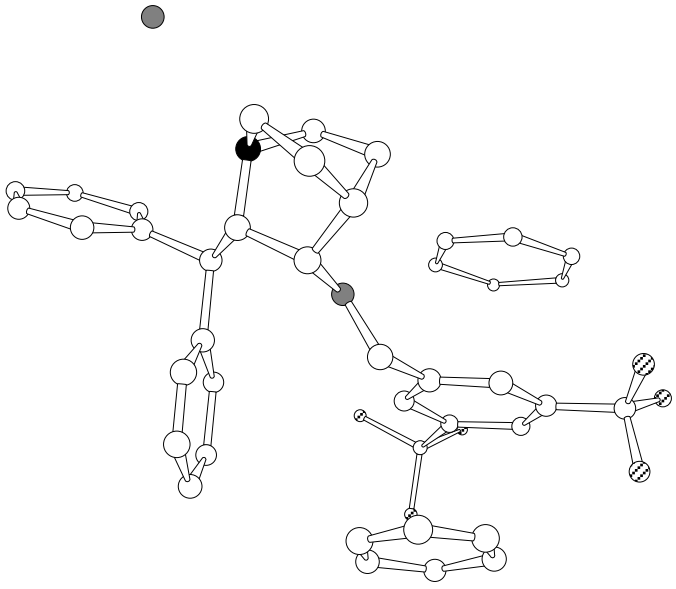


FIGURE 6. Parallel (face-to-face) aromatic stacks involving the benzylether group of NK_1 antagonist L-709,210 and benzene rings resulting from XED dockings (D=3).

employed (Fig. 5). The latter reflects the reported sensitivity of binary complex geometries to the dielectric constant when using XEDs.[1a,c] XEDs have originally been parametrized to reproduce experimentally or quantum-chemically determined geometries of binary complexes using a dielectric constant D=1.[1c] It can, however, be argued whether this is correct since the interaction geometries employed in the parametrization did not only stem from the gas phase, but also from solution and solid state results.[1c] Our results suggest that a dielectric constant of 3 may best be employed in using solid state results: aromatic stacking at the benzhydryl group of a quinuclidine antagonist as well as face-to-face aromatic stacking of

the benzylether group of L-709,210 was predicted in agreement with crystallographic results, e.g. Fig. 6 versus Fig. 7 in Chapter 2.

Favourable interaction sites for a polarizable probe are located near N2 and at the benzhydryl pocket. Sites for such a probe have been referred to as indicative for short-range surface attraction, built-up from induced electrostatics and van der Waals dispersion forces.[1b] Indeed, these 'surface points' (including the ones with interaction energies above -10.0 kcal/mol, see Methods) were found to follow the Connolly surface[25] (probe 1.4 Å) of the LEWCUL complexes (data not shown). The favourable probe positions that have been mentioned in the results were located at sites with high curvature.

Preferred intermolecular interaction geometries in small-molecule crystal structures and in protein-ligand complexes have been found to be in good agreement.[26] In agreement with the results described above XED dockings may thus be employed to study favourable interaction geometries of a ligand and functional groups in a receptor. The method should, in that respect, be further evaluated (e.g. with respect to the dielectric constant employed) and perhaps tailored with ligands 'extracted' from their macromolecular binding sites. The construction of a model for the NK_1 antagonist binding site may then also be possible. For a small ligand and a relatively small probe molecule direct dockings based on XEDs were shown to give accurate and useful results with respect to interaction geometries and a specific aromatic interaction site. Yet, for docking two larger molecules or of a ligand to a receptor docking of complementary extrema positions of the two molecules may be the method of choice. Such a method may yield the most favourable directions of approach and allows for adjustment of the conformation of the flexible ligand to the requirements of the binding site. As a result of such dockings the molecular van der Waals surfaces of the ligand and its receptor will then typically be ca. 3 Å away (i.e. separated by coinciding complementary extrema). Subsequent full relaxation, using XEDs directly, will then yield a model of the final complex.

For many receptor macromolecules that may be therapeutically interesting, but for which structural information about the ligand-receptor complex is essentially lacking, knowledge on the requirements of the ligand binding-site can only be obtained *via* series of compounds that bind to that receptor.[e.g.27] This is the case for e.g. the G-protein coupled transmembrane NK_1 receptor. When similar binding sites have been established for various ligands, or when this may be expected based on other grounds, these ligands have to reveal

similar steric and electrostatic properties in their bioactive conformations. More specifically, mimicry of the action of an endogenous ligand by an exogenous compound, such as e.g. is the case with peptidomimetics,[28] requires that the molecular 'fingerprints' of both ligands are similar. Vinter's methods may be particularly useful when, e.g. by comparison of a series of ligands, a common 'fingerprint' in terms of potential-extrema can be deduced. This 'fingerprint', which is actually a pharmacophore of interaction sites, may then be deconvoluted into a molecular framework to be embedded in novel ligands.

In conclusion, the use of XEDs as a description of the electron distribution in molecules, in combination with the calculation of potential extrema and dockings of probe molecules of a similar or smaller size appears to be a promising approach in the study of ligand-receptor interactions. The interpretation of structure-activity relationships and the study of molecular mimicry might be given a basis in terms of intermolecular interactions instead of in terms of the molecular framework, when using this approach.

ACKNOWLEDGEMENT

Preliminary studies by Irma Rigter and Muck van Groningen on the use of multi-centred point charge models on NK_1 receptor ligands and helpful discussions with Dr J. G Vinter are gratefully acknowledged.

REFERENCES

1. a) Vinter, J. G., *J. Comput.-Aided Mol. Design*, **1996**, *10*, 417; b) Vinter, J. G. and Trollope, K. I., *J. Comput.-Aided Mol. Design*, **1995**, *9*, 297; c) Vinter, J. G., *J. Comput.-Aided Mol. Design*, **1994**, 8, 653.
2. Rauhut, G. and Clarke, T., *J. Comp. Chem.*, **1993**, *14*, 503.
3. Vinter J. G and Trollope, K. I., manuscript in preparation.
4. Davies, A., Warrington, B. H. and Vinter J. G., *J. Comput.-Aided Mol. Design*, **1987**, 1, 97.
5. Rigter, I. M., Van Groningen, M., Boks, G. J., Tollenaere, J. P., Trollope, K. I. and Vinter, J. G., *Pharmacy World & Science*, **1994**, *16*, Suppl. H15.
6. Apaya, R. P., Lucchese, B., Price, S. L., and Vinter J. G., *J. Comput.-Aided Mol. Design*, **1995**, 9, 33.
7. Van der Wenden, E. M., Price, S. L., Apaya, R. P., IJzerman, A. P. and Soudijn, W., *J. Comput.-Aided Mol. Design*, **1995**, 9, 44.
8. Hunter, C A., Sanders, J. K. M., *J. Am. Chem. Soc.*, **1990**, *112*, 5525.

9. Kalindjian, S. B., Buck, I. M., Davies, J. M. R., Dunstone, D. J., Hudson, M. L., Low, C. M. R., McDonald, I. M., Pether, M. J., Steel, I. M., Tozer, M. J. and Vinter, J. G., *J. Med. Chem.*, **1996**, *39*, 1806.
10. Boks, G. J., Tollenaere J. P. and Kroon, J., *Bioorg. Med. Chem.*, **1997**, *5*, 535; and Chapter 2 of this thesis.
11. Atkins, P. W., *Physical Chemistry*, 4th ed., Oxford University Press, Oxford, 1990.
12. Tripos Associates, 1699 S. Hanley Road, St. Louis, Missouri 63144-2913 USA, 1996.
13. Jeffrey, G. A. and Lewis, L., *Carbohydr. Res.*, **1978**, *60*, 179; Allen, F. H., *Acta Cryst.*, **1986**, *B42*, 521.
14. Vinter, J. G., Davis, A. and Saunders, M. R., *J. Comput.-Aided Mol. Design*, **1987**, 1, 31.
15. Morley, S. D., Abraham, R. J., Haworth, I. S., Jackson, D. E., Saunders, M. R. and Vinter, J. G., *J. Comput.-Aided Mol. Design*, **1991**, 475.
16. Lowe, J. A., III, Drozda, S. E., McLean, S., Crawford, R. T., Bryce, D. K., Bordner, J. *Bioorg. Med. Chem. Lett.*, **1994**, *4*, 1153.
17. Steiner, Th., Starikov, E. B., Amado, A. M. and Texeira-Dias, J. J. C., *J. Chem. Soc., Perkin Trans.* **1995**, *2*, 1321.
18. Taylor, R., Kennard, O., *J. Am. Chem. Soc*., **1982**, *104*, 5063.
19. Bondi, A., *J. Phys. Chem*., **1964**, *68*, 441.
20. Verdonk, M. L., Boks, G. J., Kooijman, H., Kanters, J., Kroon, J., *J. Comput.-Aided Mol. Design* , **1993**, *7*, 173.
21. Swain, C. J., Seward, E. M., Cascieri, M. A., Fong, T. M., Herbert, R., MacIntyre, D. E., Merchant, K. J., Owen, S. N., Owens, A. P., Sabin, V., Teall, M., VanNiel, M. B., Williams, B. J., Sadowski, S., Strader, C. D., Ball, R. G., Baker, R., *J. Med. Chem.*, **1995**, *38*, 4793.
22. Dougherty, D. A., *Science*, **1996**, *271*, 163.
23. Desiraju, G. R., *Acc. Chem. Res*., **1996**, *29*, 441; Steiner, T., *Cryst. Rev*., **1996**, *6*, 1.
24. Van Mourik, T, *Correlated ab initio calculations on weakly bonded systems*, thesis, Utrecht University, 1994.
25. Connolly, M. I., *Science*, **1983**, *221*, 709.
26. Klebe, G. *J. Mol. Biol*., **1994**, *237*, 212.
27. Gether, U., Nilsson, L., Lowe, J. A., III, Schwartz, T. W., *J. Biol. Chem.*, **1994**, *269*, 23959.
28. A. Giannis and T. Kolter, *Angew. Chem. Int. Ed. Engl*., **1993**, *32*, 1244.
29. Interaction energies in the XED dockings are given with respect to the benzene ring with the highest (negative) interaction energy in each run. This is a rough approximation to the zero interaction energy at infinite separation.

Chapter 4

Crystallization of Substance P and its Retropeptoid Analogue

The undecapeptide substance P (SP) and its biologically active retropeptoid peptidomimetic were subject of crystallization experiments. Crystallization of SP was mainly prevented by the formation of amorphous precipitates in the crystallization drops. This is probably linked to the amphiphilic character of the peptide. This unfavourable behaviour could be influenced by the addition of acetonitrile or pyridine to the drops, but could so far not be prevented. The gel-like lumpy precipitates encountered in a number of the SP crystallization experiments might be the best starting point for further experiments.

Crystallization of peptides with a size comparable to that of the undecapeptide SP (H-RPKPQQFFGLM-NH_2) and its retropeptoid (Ac-nMnLGnFnFnQnQPnKPnR-NH_2)[#] (Fig. 1) has thus far only succeeded in the case of cyclic peptides, or acyclic peptides that specifically contain α,α-dialkylated amino-acid residues. This success is probably related to the considerably reduced conformational flexibility of these compounds.[1] The largest acyclic peptide containing only monoalkylated α-amino acids that, to our knowledge, successfully has been crystallized is the δ-opioid agonist hexapeptide Tyr-D-Thr-Gly-Phe-Leu-Thr.[2] In addition, only one peptoid peptidomimetic crystal structure, a peptoid analogue of the dipeptide artificial sweetener aspartame, has been reported.[3] Although SP and retropeptoid crystal structures would most likely not reveal their bioactive conformations,[4,5] they would contribute to the fundamental understanding of their energetically accessible conformations and their intermolecular interactions in a highly structured molecular environment. Successful crystallization of SP and its retropeptoid would, in addition, open promising perspectives for the crystallization of many other, highly flexible biologically active peptides and peptidomimetics.

[#] For the use of this nomenclature see Appendix A.

The chances for success of crystallizing SP *as a racemic mixture* might be larger than crystallizing a single stereoisomer. One of the indications for this is that the non-enantiomorphic space groups that are in principle available for crystal structures, where in case of chiral compounds *both* enantiomers are present in the structure, outnumber the enantiomorphic space groups (165 non-enantiomorphic versus 65 enantiomorphic).[6] However, in the crystal structures known to date molecules have been found to preferentially crystallize in only a limited number of space groups. Yet, the majority of these space groups are non-enantiomorphic.[6,7] This might be due to more favourable packing arrangements in this class of space groups than in the enantiomorphic ones.[6] Therefore, both the natural stereoisomer of SP as well as its racemate were included in our work.

Substance P, $C_{63}H_{98}N_{18}O_{13}S$, M_r 1347.8, as the acetate salt was purchased from Novabiochem. Purity according to HPLC was 98.1%, which was considered sufficient for crystallization. The constituents of the substance P racemate, *N*-acetyl substance P (L-AcSP) and the corresponding inverso *N*-acetyl substance P (D-AcSP) as the trifluoroacetate salts, have been synthesized using a solid-phase strategy using Fmoc-protected amino acids.[8] Compounds were pure on analytical HPLC. *N*-acetyl substance P was characterized by FAB mass spectrometry. $[M+H]^+ = m/z$ 1389.7. *N*-Acetyl retropeptoid substance P (trifluoro-acetate salt) was also synthesized on solid-phase, using Fmoc protected *N*-substituted glycine derivatives[8] and was pure on analytical HPLC. The covalent structure was ascertained by FAB mass spectrometry, $[M+H]^+ = m/z$ 1389.7. Compounds were stored at -20 ˚C mainly to prevent degeneration of the peptides.[9]

The crystallization strategy was based on the use of a screen of macromolecular crystallization conditions, employed in a hanging drop-vapour diffusion set-up (Fig. 2).[10] A drop of 2 µl peptide or peptoid solution (10 mg/ml) was placed on a siliconized glass coverslip with dimensions 2x2 cm. This drop was mixed with 2 µl of a precipitant solution from the crystallization screen and placed up-side-down (drop down) over a crystallization well containing 750 µl of the precipitant solution (Fig. 2) and was allowed to equilibrate. The contacts between coverslip and well had been made air-tight with grease. This approach has the advantage over conventional techniques for crystallizing small molecules,[11] that many more conditions can readily be evaluated and easily controlled. Furthermore, only relatively small amounts (2 to 10 mg) of peptide or peptoid are required for evaluating several hundreds of different conditions; typically 20 µg of sample per experiment is required. The method is in practice, however, restricted to crystallization from aqueous solution. We basically

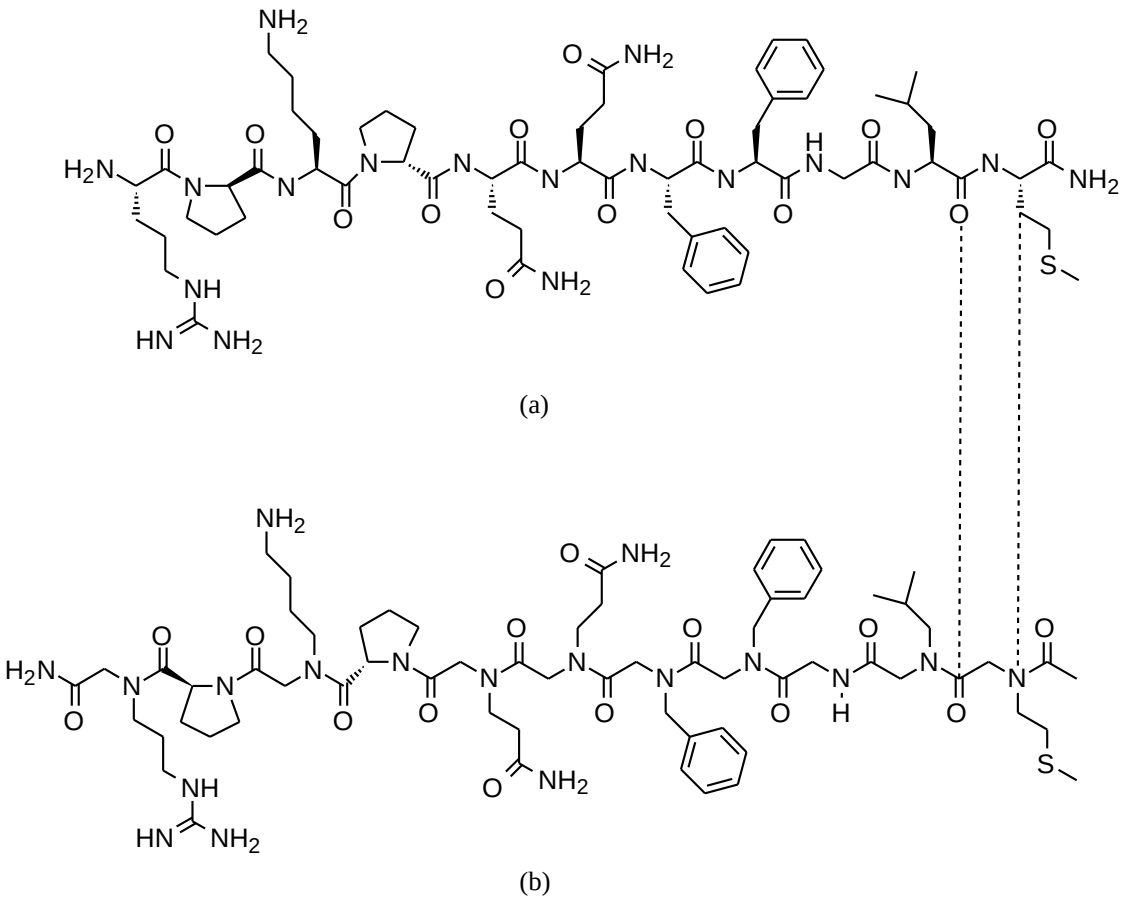


FIGURE 1. Structural formula of (a) substance P (SP) and (b) its retropeptoid peptidomimetic. The compounds are full agonists at the murine NK_1 receptor (see Chapter 1). In addition to the similar topological positions of the side chains in peptide and peptoids, in *retro*peptoids also the backbone carbonyl groups are found in equivalent topological positions as in the parent peptide.

employed a modified version of the ‘sparse matrix sampling solutions’, containing 54 different conditions (‘magic 50’).[12] The precipitant solutions in the crystallization screen are comprised of a large range of buffer (pH range 4.5 - 8.9), salt and precipitant (e.g. various PEGs (polyethylene glycol), 2-methyl-2,4-pentanediol (MPD), isopropanol, 1,2-hexanediol, salt) concentrations. Crystallization experiments were performed at 277 K and 293 K.

The precipitates in the crystallization drops for SP (10 mg/ml) that appeared within a week in the 'magic 50' crystallization screen both at 277 K and 293 K could roughly be classified into three groups. In addition to drops that remained clear we observed (a) gel-like

lumps (b) oil-like drops dispersed in solution either light or dark coloured, and (c) untransparent white-brownish precipitate. These precipitates could be characterized as amorphous precipitates (a and c) or phase-separations (b). Crystals were, however, not observed. The gel-like lumps seem to be associated with the presence of PEG (i.e. 30% (w/v) PEG3k, 30% (w/v) PEG 400, 16% (w/v) PEG20k), 30 % (v/v) isopropanol or sodium acetate (1 or 1.4 M) in the precipitant solutions whereas the white-brownish precipitate occurs at high salt concentrations (1.4 M sodium citrate, 2M ammonium sulfate and 4M sodium formate).¶ The oil-like drops were obtained under various conditions without a clear common factor. Different precipitates sometimes occurred at different temperatures, but with otherwise identical conditions. Many precipitates, however, were largely unaffected by the different temperatures. The gel-like lumps and the white-brownish precipitate gradually (several months) redissolved when these drops were allowed to equilibrate over a crystallization well containing only water. The oil-like drops, on the other hand, remained unchanged. The formation of the former two precipitate forms was therefore considered to be reversible.

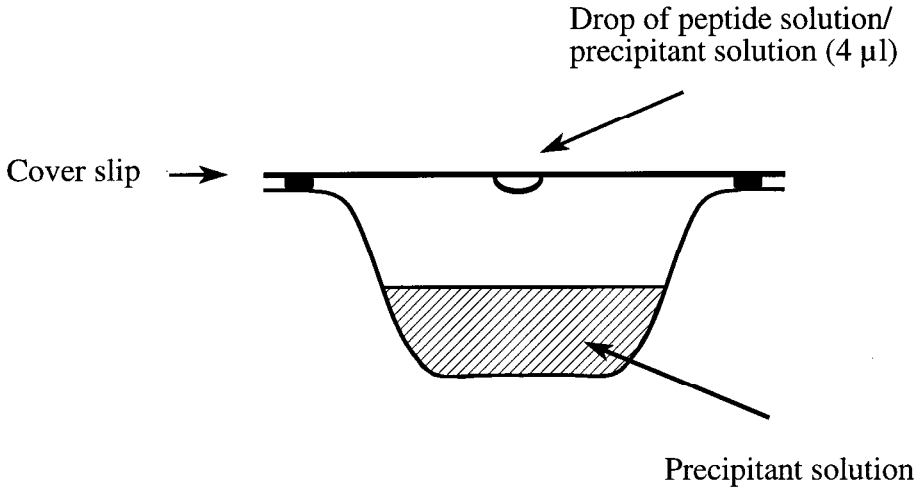


FIGURE 2. Hanging drop-vapour diffusion set-up. The initial precipitant concentration in the crystallization drop is half the concentration in the reservoir. The closed system reaches equilibrium by a net diffusion of water (or reversed diffusion of other volatile compounds) from the drop to the reservoir. Supersaturation can so be reached in the drop and the peptide or peptoid precipitates.

¶ PEG is known for possibly resulting in gel-like precipitates under conditions with high PEG concentrations, yet also some non-PEG containing solutions yielded gel-like lumps in our experiments.

Precipitation and aggregation behaviour of SP in aqueous solution under the influence of varying pH, salt- and peptide concentrations has been described before.[13-16] This behaviour is probably related to the amphiphilic nature of the peptide (Fig. 3). Organic solvents such as acetonitrile and pyridine were found to inhibit or reverse this behaviour[16] and were as such employed in our crystallization experiments. This was performed by adding 1 µl acetonitrile or pyridine to a 1µl drop of peptide solution, before mixing with the precipitant solution. A selection of precipitant solutions from the 'magic 50' screen was employed in this procedure. A clear effect was observed on the morphology of the precipitates, when compared to the absence of the additives, but further application of these solvents was hampered by their volatility and was not pursued.

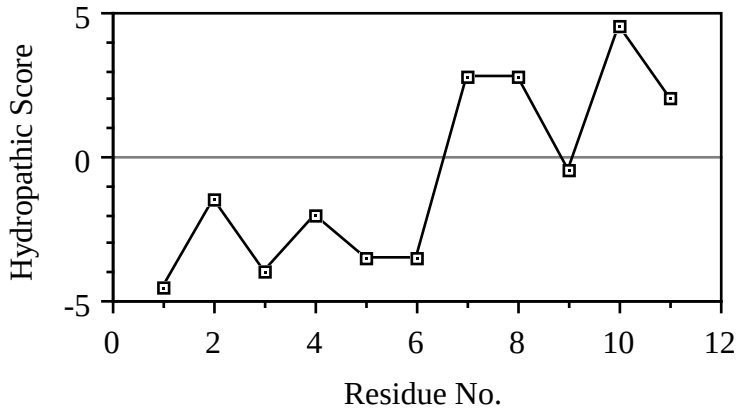


FIGURE 3. Hydropathy profile of substance P based on the Kyte and Doolittle hydropathy scale for amino acids.[16] Substance P is clearly amphiphilic: the C-terminal pentapeptide is largly hydrophobic, whereas the N-terminal hexapeptide is largely hydrophilic on this scale. (After: Poujade et al.[16])

A considerably lower total peptide concentration (2 mg/ml) was employed for the ‘magic 50’ experiments with the racemate of SP. The precipitates observed resemble the amorphous precipitates of the enantiomerically pure SP, but were less extensive. Crystals have not been observed so far. The crystallization experiments with the SP retropeptoid (10 mg/ml) did not lead to initial crystallization conditions so far. In fact, most of the crystallization drops remained essentially clear after six months. The aqueous solubility of the SP retropeptoid is markedly lower than the solubility of SP; SP dissolved readily in water up to a concentration of 10 mg/ml, whereas the retropeptoid yielded a suspension at that

concentration. This, in light of the results, suggests that the effective peptoid concentration in the drops was considerably lower than 10 mg/ml. Remarkably, both the 2 µl drops of SP and retropeptoid solution to which precipitant solution was added, were slightly coloured brown at ambient temperature, but were colourless at 277 K.

There are several reasons to assume that the conditions leading to gel-like lumps in the crystallization attempts with SP that cannot be attributed to PEG aggregation, might be starting points for successful crystallizations. It is quite common in macromolecular crystallizations that amorphous precipitates (as they are generally called, including the gel-like lumps) occur.[17] One of the models that describes macromolecular crystal growth involves an initial non-crystalline crystallization nucleus that results in a crystal lattice under the appropriate conditions for growth, or to the kinetically favoured amorphous precipitates.[18] Also, in some cases a clear distinction cannot *a priori* be made between an amorphous structure and a regular crystal lattice; a diffraction experiment is required for obtaining conclusive results.[18] A slight change in conditions can, in addition, make the difference between an amorphous or crystalline precipitate.[17] Furthermore, crystal growth from amorphous precipitates has been observed before (Van den Elsen, personal communication). SP (eleven residues) can, however, not be considered a macromolecule, as is already reflected in the exceptionally high conformational flexibility of the backbone in solution[19] in comparison to proteins. Yet, the gel-like precipitates are what comes closest to a somewhat ordered structure of SP. These precipitates were considered insufficiently solid so that preliminary X-ray analysis was not considered useful. Anyway, in the case of the amorphous SP precipitates, supersaturation probably proceeded too fast.[17] Since also clear drops remained present in the 'magic 50' screen after several months, there is ample opportunity for assessing crystallization conditions in between these two extremes.

Can we ever expect to obtain crystals of SP suitable for X-ray analysis? The hydrophobic clustering and aggregation behaviour of SP,[16] that is probably related to its amphiphilic nature, has precluded crystallization thus far. Ample opportunities, however, exist to further explore the precipitation behaviour of SP, whether as a single stereoisomer or as a racemic mixture, starting at the conditions of the 'magic 50' screen. Also, the biologically active C-terminal hexapeptoid analogues of SP (see Chapter 1) may be employed in the crystallization experiments; these peptoids are less flexible and are largely devoid of the amphiphilic character of SP. For obtaining fundamental knowledge on the structure,

interactions and conformational behaviour of medium sized peptides and peptidomimetics (six to thirty residues) under the influence of different environments[20] (e.g. by trying to crystallize different polymorphs) a successful crystallization strategy would present a tremendous break-through. When, on the other hand, biological activity of the compounds or other processes involving molecular recognition between a peptide or peptidomimetic and a (macromolecular) receptor are of main interest, crystallization of the receptor-ligand complex should be the method of choice.

ACKNOWLEDGEMENT

GB wishes to thank Jolanda Hurenkamp for introducing him to the practical part of protein crystallization. Arie Schouten and Dr Piet Gros are gratefully acknowledged for helpful discussions and for critically reading the manuscript.

REFERENCES

1. Karle, I. L., *Acta Cryst*., **1992**, *B48*, 341.
2. Flippen-Anderson, J. L., Deschamps, J. R., Ward, K. B., George, C. and Houghten, R., *Int. J. Peptide Protein Res*., **1994**, *44*, 97.
3. Boks, G. J., Boer, D. R., Kruijtzer, J. A. W., Liskamp, R. M. J., Tollenaere, J. P., Kroon, J. and Schouten, A., *in preparation*, **1997**.
4. Nicklaus, M. C., Wang, S., Driscoll, J. S. and Milne, G. W. A., *Bioorg. Med. Chem*., **1995**, *3*, 411.
5. Deschamps, J. R., George, C. and Flippen-Anderson, J. L., *Biopolymers*, **1996**, *40*, 121.
6. Brock, C. P., Schweizer, W. B. and Dunitz, J. D., *J. Am. Chem. Soc.*, **1991**, *113*, 9811.
7. (a) Mighell, A. D., Himes, V. L., Rodgers, J. R., *Acta Cryst*., **1983**, *A39*, 737; (b) Donohue, J., *Acta Cryst*., **1985**, *A41*, 203
8. Kruijtzer, J. A. W., *Synthesis of Peptoid Peptidomimetics*, thesis, Utrecht University, 1996.
9. Kertscher., U., Bienert, M., Krause, E., Sepetov, N. and Mehlis., B., *Int. J. Peptide Protein Res*., **1993**, *41*, 207.
10. Blundell, T. L. and Johnson, L. N., Protein Crystallography, Academic Press, New York **1976**.
11. Van der Sluis, P., *Single Crystals and X-Ray Structure Determination*, thesis, Utrecht University, 1989.
12. Jancarik, J. and Kim, S.-H., *J. Appl. Cryst*., **1991**, *24*, 409.
13. Rueger, M., Bienert, M. and Mehlis, B., *Biopolymers*, **1984**, *23*, 747.

14. Mehlis, B., Rueger, M., Becker,. M., Bienert, M., Niedrich, H. and Oehme., P., *Int. J. Peptide Prot. Res*., **1980**, *15*, 20.
15. Mehlis, B., Boehm, S., Becker, M. and Bienert, M., *Biochem. Biophys. Res. Commun.*, **1975**, *66*, 1447.
16. Poujade, C., Lavielle, S., Chassaing, G. and Marquet, A., *Biochem. Biophys. Res. Commun*., **1983**, *114*, 1109.
17. McPherson, A., *Preparation and Analysis of Protein Crystals,* Krieger Publishing Company, Malabar 1982.
18. McPherson, A., Malkin, A. J. and Kuznetsov., Y. G., *Structure*, **1995**, *3*, 759.
19. Chassaing, G., Convert, O. and Lavielle, S., *Eur. J. Biochem*., **1986**, *154*, 77.
20. Montcalm, T., Cui., W., Zhao., H., Guarnieri., F. and Wilson, S. R., *J. Mol. Struc*., **1994**, *308*, 37.

Chapter 5

Geometrical Features of Tertiary and Secondary Amide Bonds in Crystal Structures: Implications for Peptoids

ABSTRACT

According to *in vacuo* quantum chemical calculations tertiary amide bonds prefer a markedly non-planar geometry. Recently, gas electron-diffraction (GED) experiments revealed,[19] however, that the out-of-plane potential for pyramidalization and inversion at the tertiary amide nitrogen is very shallow and that no clear-cut conclusions can be drawn regarding the (non)planarity of tertiary amides in the gas phase. In the solid state a considerable out-of-plane deformation is possible, but the tertiary amide bonds prefer to be planar. It is likely that this is due to the presence of a structured environment. The trigonal amide nitrogen coordination geometry is considerably more flexible for out-of-plane movements than the carbonyl carbon coordination geometry as expressed by the ratio of the standard deviations of the corresponding out-of-plane parameter distributions. In protein binding sites pyramidalization at the amide nitrogen in tertiary amide bonds occurs to a similar degree as in the solid state. In conclusion, pyramidalization at the tertiary amide nitrogen atom ranges from about one-third in the solid state and in protein binding sites to one half in the gas phase compared to that of a full sp3 coordination geometry, expressed as the out-of-plane parameter χ_N. In *trans* amide bonds along the main chain, the side chain in peptoid monomers shows the largest out-of-plane position. In *cis* amide bonds, on the other hand, the out-of-plane position of the C_α atom is likely to be mainly 'responsible' for any pyramidalization at the amide nitrogen atom. Pyramidalization at the tertiary amide nitrogen seems to be an additional possibility for peptoids to accommodate their conformation in order to optimize complementarity with the protein binding site.

INTRODUCTION

Structurally, peptoids can be regarded as chains of *N*-alkylated glycine, proline and glycine residues connected by tertiary or secondary amide bonds (Fig. 1). In comparison with peptides, the amino-acid side chains have been shifted from the chiral C_α carbon atom to the amide nitrogen atom. For proline and glycine residues this particular modification leaves them unaltered. In addition to the loss of the chiral carbon atom in the non-proline peptoid residues, when compared to α-amino acids, a larger conformational flexibility about the C_α methylene groups can be expected. The amide bonds remain the most rigid parts of the peptoid chain, yet have lost their potential N-H hydrogen bond donors (the glycine residues excluded). The location of the side chains will, in analogy with peptides, be completely determined by the conformation of the peptoid backbone. In this respect it is, however, important to note that *cis/trans* isomerization of the tertiary amide bond in peptoids has, in contrast to the situation in peptides, been observed in different solvents at ambient temperature.[1,2]

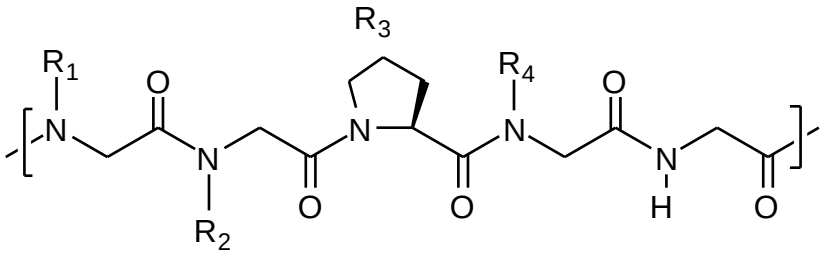


FIGURE 1. Generalized peptoid backbone structure, also containing a proline and a glycine residue.

Amide bonds thus play an important role in the structure of peptoids. Although secondary amide bonds are generally found to be planar, out-of-plane deformations can occur in ring systems[3] or under the influence of approaching nucleophiles in the solid state[4,5]. The coordination geometry of the secondary amide nitrogen was found to be several times more flexible than the carbonyl carbon coordination geometry[3]. This behaviour can generally be rationalized in terms of the electronic structure of the amide bond (Fig. 2). The amide bond resonance forms can be considered as limiting situations of the variation in the electron distribution.[6] Changing a secondary into a tertiary amide bond reduces the partial double bond character of the central C-N bond and might result in a larger flexibility at the amide

nitrogen. Furthermore, depending on the specific environment in a protein binding site, pyramidalization at the amide nitrogen atom may occur. In order to quantify this effect, we determined the range of the out-of-plane deformation of the coordination geometry at the amide nitrogen and carbonyl carbon atoms in tertiary and secondary amide bonds in a collection of small molecule crystal structures, *viz.* the Cambridge Structural Database.[7] This may give insight in the degree of pyramidalization that can occur in a structured (solid state) environment as well as in molecular structures that are generally believed to be in a low-energy state.[8] A similar analysis was performed for the amide bonds of sarcosine (*N*-methylglycine), proline and glycine residues in protein binding sites in crystallographically determined protein complexes. Results have been compared to *in vacuo* amide bond geometries of a model peptoid in quantum chemical calculations at the MP2/6-31G* level[9] and to empirically calculated amide bond geometries using the CFF91 force field.[10]

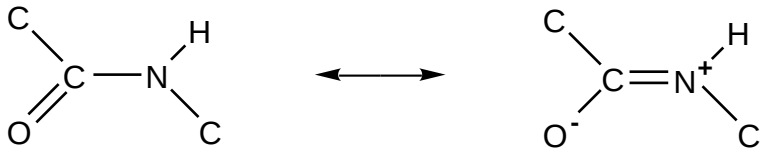


FIGURE 2. Resonance forms of a secondary amide group

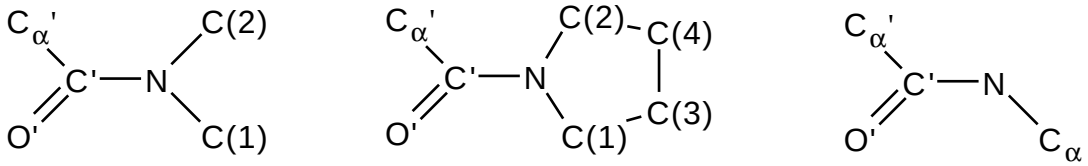
METHODS AND MATERIALS

Substructure Searches in the Cambridge Structural Database

Substructure searches were performed using QUEST 2.17 and the results were analysed with VISTA 2.3, both of which are part of the Cambridge Structural Database System.[7] The CSD was searched for acyclic tertiary amides (**1**), tertiary amides in which the amide nitrogen is part of a pyrrolidine ring such as in proline (**2**), and secondary amides (**3**) (Fig. 3). C_{α}', C_{α}, C(1), C(2), C(3) and C(4) are sp3 hybridized carbon atoms. To prevent ambiguity in assigning C(1) and C(2) based on topology, C(1) is the carbon atom more or less eclipsed with respect to O' (O'-C'-N-C(1) within -90 and 90˚). The atom labelling employed is based on the IUPAC-IUB nomenclature for peptides.[11] A prime indicates that the atom belongs to a preceding residue in the chain. No additional covalent bonds were allowed between atoms beside the ones indicated in the structural diagrams (Fig. 3).

Hydrogen bonds accepted by the carbonyl oxygen atoms were permitted, but covalent coordination by e.g. metal atoms was excluded.

Structures with a conventional R-factor ≤ 0.10 and a covalent bond tolerance of 0.4 Å were included in the search. Structures that were error-flagged, contained disorder or were part of polymers were excluded. Duplicate structures were also discarded. Fragments that were outliers in a bond length or a bond angle value by more than four times the standard deviation from the mean value of the corresponding parameter in the set were discarded. The geometry of the achiral fragments was determined in an asymmetric part of the fragment's conformational space (see the Experimental section).[12]



1 **2** **3**

FIGURE 3. Substructures characteristic for (**1**) acyclic tertiary amide bonds, (**2**) tertiary amides of which the amide nitrogen is part of a pyrrolidine ring such as in proline and (**3**) secondary amides (**3**). In analogy with the atom labelling scheme adopted for peptides,[11] atoms belonging to another residue are indicated with a prime.

Amide Residues in Protein Binding Sites

Crystallographically determined protein-ligand complexes containing sarcosine (*N*-methylglycine), proline or glycine as part of the ligands were retrieved from the Brookhaven Protein Data Bank[13] (PDB, November 1996). The search was restricted to well resolved structures with a high resolution (≤ 2.0 Å) (≤ 2.5 Å for sarcosine) and an R-factor ≤ 0.20. In addition, atomic positions for all atoms of the residues of interest, as well as their N-terminal amide bonds should have been explicitly determined. Proline and glycine residues at the N- or C-terminus of a ligand were excluded. In contrast to the substructure searches in the CSD, where the fragment conformation was determined in an asymmetric part of the conformational space, here the residue conformations were evaluated as present in the protein binding sites.

RESULTS

Small-Molecule Crystal Structures

Dunitz and Winkler[3] defined the out-of-plane bending parameters χ_N and χ_C and a twisting parameter τ, which completely describe the out-of-plane deformation of an amide group. The parameters are defined using the amide bond torsion angles. Definitions are given in Fig. 4. χ_N is the clockwise rotation of C(2) to C(1) to obtain a mutually antiparallel orientation of C(2) and C(1), when viewed along the direction C'-N (Fig. 4). χ_C is the *counter* clockwise rotation of C_α' to O' when viewed along C'-N to obtain a perfect antiparallel orientation with respect to O'. The twisting parameter τ is the dihedral angle between the normals of the planes through C_α', C', O' and C(2), N, C(1), originating at C' and N, respectively, and determined along the C'-N bond (Fig. 4). For *trans* and *cis* amide bonds along the main chain τ values are approximately 180 and 0˚, respectively. C(1) is in the latter case replaced by C_α in the parameter definitions. Distributions of the three out-of-plane parameters for tertiary and secondary amide bonds fragments in small molecule crystal structures are given in Figs 5 (a)-(g). Due to symmetry, the right- and left-hand side of every distribution are each others mirror image (see the Experimental section). Average solid-state bond lengths and bond angles for the amide bonds in Fig. 3 as determined from our sets are given in Appendix B.

In strict analogy with the common definition for the variation of a normal distribution we calculated a measure of variation for each distribution given in Fig. 5 in order to quantify the relative measure of the out-of-plane movements. This variation is defined as the sum of the squared differences of each out-of-plane parameter value to its "average value" divided by the number of degrees of freedom. The "average values" of the distributions in Fig. 5, which are either 0 or 180˚, do not have a physical meaning (see the Experimental section for details). Only the highest occupied classes have a physical meaning: they represent the preferential parameter values. The preferential values, the square roots of the variations, i.e. the standard deviations, and the extreme values of the distributions are given in Table 1.

Protein-Ligand Complexes

For comparison, values for the amide out-of-plane parameters were also evaluated in sarcosine, proline and glycine residues that are part of ligands in well resolved

crystallographically determined protein complexes. The minimum and maximum parameter values are given in Table 2. Considerable redundancy is present in the sets for all three residue types with respect to the protein binding sites, e.g. 28 out of the 30 tabulated sarcosine residues are bound at the S3 binding site in various cyclophilins complexed with cyclosporin A (see Table 1, Chapter 6). Therefore, the extreme parameter values rather than the complete distributions are given.

In the structure of carbamoyl sarcosine complexed with creatine amidinohydrolase (creatinase, PDB entry code 1CHM) the amide bond at the N-terminal side of the sarcosine residue is part of a urea group;[14] the conjugated system thus is not restricted to the amide bond, as is the case for all other fragments. The tertiary amide nitrogen reveals a tetrahedral coordination geometry (χ_N= 60 (1CHM:CMS A404), χ_N= 64˚ (1CHM:CMS B404)) that is stabilized by an extensive hydrogen bond network in which the tertiary amide nitrogen functions as a hydrogen bond acceptor for a nearby His residue.[14] This entry was not included in Table 2.

TABLE 1. Characteristics of the out-of-plane parameter distributions as given in Fig. 5 for tertiary (**1** acyclic, **2** cyclic) and secondary (**3**) amide bonds

	1 n=104		**2** n=108		**3** n=46	
Parameter	Preferential class (˚)	σ (˚)	Preferential class (˚)	σ (˚)	Preferential class (˚)	σ (˚)
χ_C	0	2.9	0	4.1	0	3.4
χ_N	0	8.5	-2.5, 2.5	11.2	-	-
τ	180	5.5	180	6.4	-	-
	Min./Max.[a] (˚)		Min./Max. (˚)		Min./Max. (˚)	
$\lvert\chi_C\rvert$	5.3		10.4		7.2	
$\lvert\chi_N\rvert$	21.0		20.1		-	
$\lvert\tau\rvert$	169.0		168.9		-	

The amide bond fragments are depicted in Fig. 3. a) Due to the symmetry of the distributions only the absolute value of the extreme values is given.

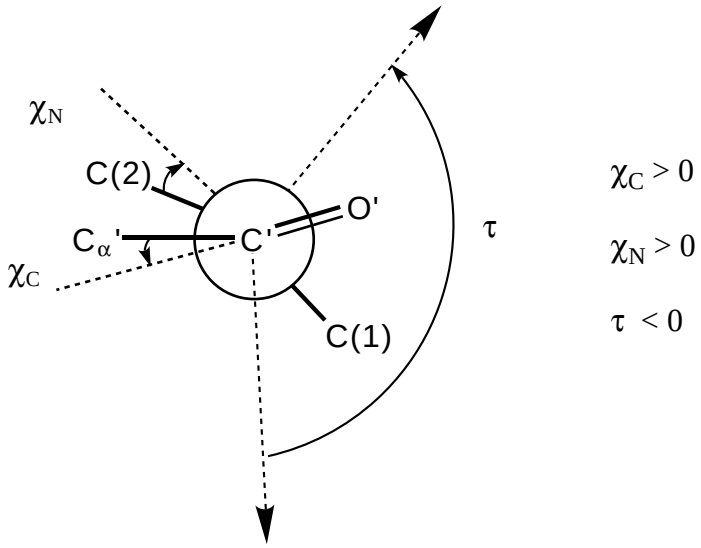


FIGURE 4. Newman projection of a tertiary amide bond indicating the out-of-plane bending parameters χ_C and χ_N and the twisting parameter τ for tertiary amide bonds. Torsion angles: $\omega1=\omega(C_\alpha'\text{-}C'\text{-}N\text{-}C(1))$, $\omega2=\omega(O'\text{-}C'\text{-}N\text{-}C(2))$, $\omega3=\omega(O'\text{-}C'\text{-}N\text{-}C(1))$, $\omega4=\omega(C_\alpha'C'\text{-}N\text{-}C(2))$. Out-of-plane bending parameters: $\chi_C=(\omega1-\omega3+\pi) \bmod 2\pi$, $\chi_N=(\omega2-\omega3+\pi) \bmod 2\pi$. Twisting parameter: $\tau=(\omega1+\omega2)/2 \bmod 2\pi$. In the expression for τ, the torsion angle values should here be taken in the range 0-360˚ or from -180 to 180˚ in case of a *trans* or *cis* amide bond, respectively The notation *(argument) mod 2π* means that the resulting argument value is brought to an interval within -180 and 180˚, in line with the normal torsion angle definitions.

TABLE 2. Variation in the planarity of tertiary and secondary amide bonds in sarcosine, proline and glycine residues in ligands in well resolved protein complexes

	Sar n=30		Pro n=73		Gly n=19	
Parameter	Min. (˚)	Max. (˚)	Min. (˚)	Max. (˚)	Min. (˚)	Max. (˚)
χ_C	-9	5	-7	10	-7	3
χ_N	-10	9	-23	21	-	-
τ	174	-176	158	-170	-	-

Out-of-plane parameter values that are well out of range with respect to the values in the Table are χ_N = -45 degrees for Pro449 in subtilisin Carlsberg complexed with an inhibitor (1SCN: Pro449),[15] and χ_C = -23 (1CDK: GlyI10) and χ_C =-31 degrees (1CDK: GlyJ10) in a cAMP-dependent protein kinase catalytic subunit-inhibitor complex.[16]

Gas Phase Geometries of Tertiary Amide Bonds

Recently, Moehle and Hofmann[9] published (local) minimum energy conformations of the model peptoid Ac-NAla-NMe2 (**4**) (Fig. 6) at various *ab initio* quantum chemical and (semi)empirical levels *in vacuo*. The out-of-plane parameter values calculated from their results on the *trans* amide bond conformer of **4** at the MP2/6-31G* level as well as some amide bond torsion angles are given in Table 3. Results are also given for these local minimum energy conformations when recalculated (re-optimized) using the CFF91 force field in DISCOVER 2.96 (INSIGHTII 2.3.5)[17] (see the Experimental section). RMS differences for the non-hydrogen atoms of the conformations optimized at the MP2/6-31G* level and with CFF91 were 0.137 Å for conformation $C_{7\beta}$, 0.157 Å for α_D and 0.117 Å for α. See Table 3 for the main chain ϕ,ψ torsion angles in these conformations. The conformations are denoted in accordance with Moehle and Hofmann.[9] The reported[9] $C_{7\beta}$ conformation containing a *trans* amide bond was reoptimized at the HF/6-31G* level to assess the influence of electron correlation on the final amide bond geometries. *Ab initio* calculations were performed within the SPARTAN program package.[18] The out-of -plane parameter values for the N-terminal tertiary amide bond (χ_C = -1.4, χ_N =-18.4 and τ =179.6˚) are in excellent agreement with the values at the MP2/6-31G* level (Table 3). As a positive control of our calculations the final ϕ,ψ torsion angles corresponded within 0.5˚ with the values from HF/6-31G* level calculation, that have been reported previously.[9]Additional geometry optimizations were performed on the model compound *N*,*N*-dimethylacetamide ($CH_3CON(CH_3)_2$ or $AcNMe_2$) using a 6-31G* basis set, starting from a perfectly planar geometry (symmetry C_s) or with slight pyramidalization at the amide nitrogen (symmetry C_1). Due to the imposed C_s symmetry -the molecule contained a mirror plane- the amide bond remained perfectly planar, whereas a starting geometry with C_1 symmetry resulted in a non-planar amide bond with clear pyramidalization at the amide nitrogen (χ_C = 1.2, χ_N = 16.8 and τ = -176.9˚). The non-planar geometry of *N*,*N*-dimethylacetamide was only 71 cal·mol^{-1} lower in energy than the planar one, indicating an extremely flat energy profile for the out-of-plane movement at the amide nitrogen.

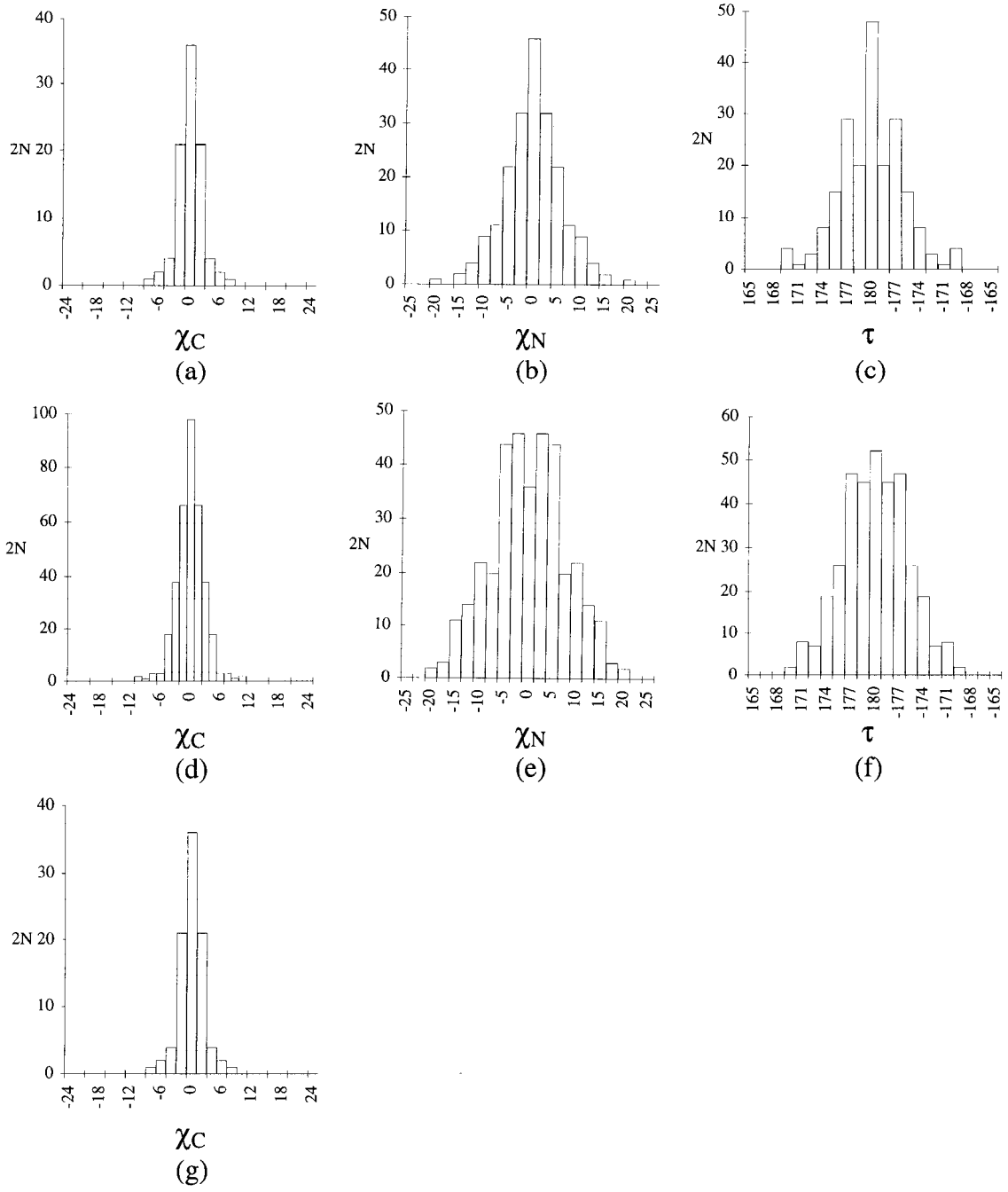


FIGURE 5. Distributions for the χ_C, χ_N and τ out-of-plane bending parameters[3] in tertiary and secondary amide bonds in small molecule crystal structures. Distributions are given for the entire conformational space, with as result that the right-hand and left-hand side of the distributions are each other's mirror image (see the Experimental section for details). Top row: acyclic tertiary amide bonds (**1**), (a) χ_C, class width 1.0˚, (b) χ_N, class width 2.5˚, (c) τ, class width 1.5˚. Middle row: cyclic tertiary amides (**2**), (d) χ_C, class width 1.5˚, (e) χ_N, class width 2.5˚, (f) τ, class width 1.5˚. Bottom row: secondary amides (**3**), (g) χ_C, class width 2.0˚.

DISCUSSION

Gas-Phase Geometries

Recently, Mack and Oberhammer have reported on the geometry of *N,N*-dimethylacetamide in a gas electron-diffraction (GED) experiment.[19] GED studies determine the average interatomic distances in a molecule. In the refinement of the gas-phase structure it appeared that both a rigid molecular model with only small-amplitude vibrations, and a dynamical model with a large amplitude vibration at the amide nitrogen atom, performed equally well in explaining the experimental observations. The rigid model resulted in a non-planar equilibrium structure[#] at the amide nitrogen, whereas the dynamical model proposed either a planar or pseudo-planar equilibrium structure with large out-of-plane vibrations of the amide nitrogen.[19] This ambiguity may be explained by the finding that the experimental potential barrier for pyramidal inversion (15 cal·mol^{-1}) is below the calculated vibrational ground-state energy level for the nitrogen out-of-plane vibration (140 cal·mol^{-1} at HF/3-21G level).[19] On the other hand, no disagreement exists with respect to the finding that the out-of-plane potential for inversion and pyramidalization at the tertiary amide nitrogen in the gas phase is very shallow. Our results substantiate these findings for the solid state and for protein binding sites. Also, in a recent high quality quantum chemical study (including MP2/6-31G* calculations) primary amino groups in nucleic acid bases, aniline, aminopyridines and aminotriazines were found to be non-planar.[20] The χ_N values that we calculated from the reported[20] 'benchmark' results (MP2/6-311G(2*df,p*) level) ranged from 22.4˚ in aminotriazine, 34.0˚ in cytosine to 56˚ in aniline. Calculated vibrational frequencies corresponded quite well with experimental values for aniline, and supported the validity of the theoretical (quantum chemical) model.[20] Experimental values for the (non)planar geometry of the amino group in these compounds are, however, still unknown. The largest out-of-plane coordination geometries for the tertiary amide bond nitrogen atom according to *in vacuo* calculations (Table 3) are located at the *lower* end of the χ_N value range in the amino compounds given above.

In view of these, as well as our own finding we hold the view that in the gas-phase the acyclic tertiary amide bond geometry may be considered to be non-planar. In this respect, we consider the local minima energy conformations for Ac-NAla-NMe2 (**4**) (Fig. 6) in the

[#] The sum of the bond angles involving the amide nitrogen correspond to $\Sigma\alpha_N$=354.1(17)˚.[19] In the solid state this would correspond to a $|\chi_N|$ value of ca. 25-30˚.

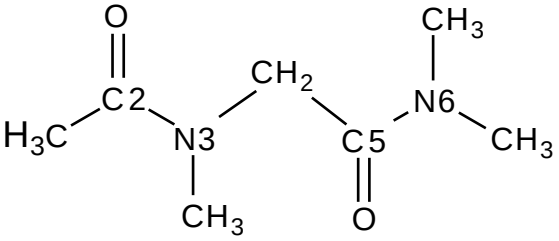


FIGURE 6. Atom labelling scheme for the model peptoid Ac-NAla-NMe_2 (**4**).[9]

TABLE 3. Out-of-plane parameter values for the tertiary amide bonds in local minimum conformations of Ac-NAla-NMe2 (**4**) at MP2/6-31G* level and recalculated using CFF91

	$C_{7\beta}$		α_D		α	
ϕ,ψ-values	(-128.2, 77.0)		(74.2, -175.6)		(-54.7, -47.2)	
Parameter	MP2/6-31G*	CFF91	MP2/6-31G*	CFF91	MP2/6-31G*	CFF91
χ_{C2} (°)	-1.8	-0.2	-2.0	-0.6	-1.7	-1.4
χ_{N3} (°)	-17.1	2.9	-30.4	-12.1	-23.9	-23.3
τ(C2-N3) (°)	180.0	177.9	176.9	-179.3	175.8	176.7
ω1(C2-N3) (°)	-172.3	nd[a]	-168.9	nd	-173.1	nd
ω4(C2-N3) (°)	-9.4	nd	-19.2	nd	-17.0	nd
χ_{C5} (°)	1.1	-1.5	2.1	3.6	4.0	3.2
χ_{N6} (°)	-11.0	-16.9	19.1	18.3	3.6	-1.2
τ(C5-N6) (°)	177.8	178.4	-177.8	-179.2	172.1	173.1
ω1(C5-N6) (°)	-176.2	nd	173.7	nd	172.2	nd
ω4(C5-N6) (°)	-7.2	nd	12.8	nd	-4.1	nd

The given torsion angles are ϕ(C2-N3-C-C5) and ψ(N3-C-C5-N6), see Figure 6. Out-of-plane parameter values have been calculated from the results reported by Moehle and Hofmann.[9] a) not determined.

potential energy hyperspace *in vacuo* as calculated using a 6-31G* basis set including electron correlation (MP2 level),[9] as an important theoretical reference point. The effect of including electron correlation via second-order Møller-Plesset perturbation theory (MP2), in comparison to the Hartree-Fock (HF) calculation, was largely reflected in the energy of the optimized structures[9] and only resulted in a minor difference in the backbone conformation as well as in the amide bond geometries (*vide infra*). The model compound *N*,*N*-dimethylacetamide ($AcNMe_2$) remained non-planar upon geometry optimization at the HF/6-31G* level. Therefore the out-of-plane geometry of the tertiary amide bonds in the optimized structures of Ac-NAla-NMe_2 (MP2/6-31G*) can probably neither be attributed to an increased number of steric interactions in the latter molecule nor to including electron

correlation. Rather, pyramidalization at the amide nitrogen seems to be an intrinsic property of tertiary amide bonds *in vacuo*, within the approximations of the theoretical quantum chemical model. In most cases pyramidalization at the amide nitrogen is considerably larger than at the carbonyl carbon atom (Table 3). In many cases the results for the out-of-plane geometries in CFF91 and MP2/6-31G* calculations correspond remarkably well (Table 3).

Solid-State Geometries

In crystal structures all three amide bond fragments (**1**, **2** and **3**) prefer a planar coordination geometry at the carbonyl carbon atom (Fig. 5(a,d,g)), as well as at the nitrogen atom in acyclic tertiary amide bonds (Fig. 5 (b)). In contrast, in proline amide bonds the distribution for the out-of-plane parameter χ_N suggests that the amide nitrogen prefers a slight (2.5 to 5˚) out-of-plane coordination geometry (Fig. 5 (e)). Whether the depression at $\chi_N = 0$ is significant can not be stated unequivocally with the present data. It is, however, plausible that the preferential geometries of the five-membered pyrrolidine ring are best accommodated

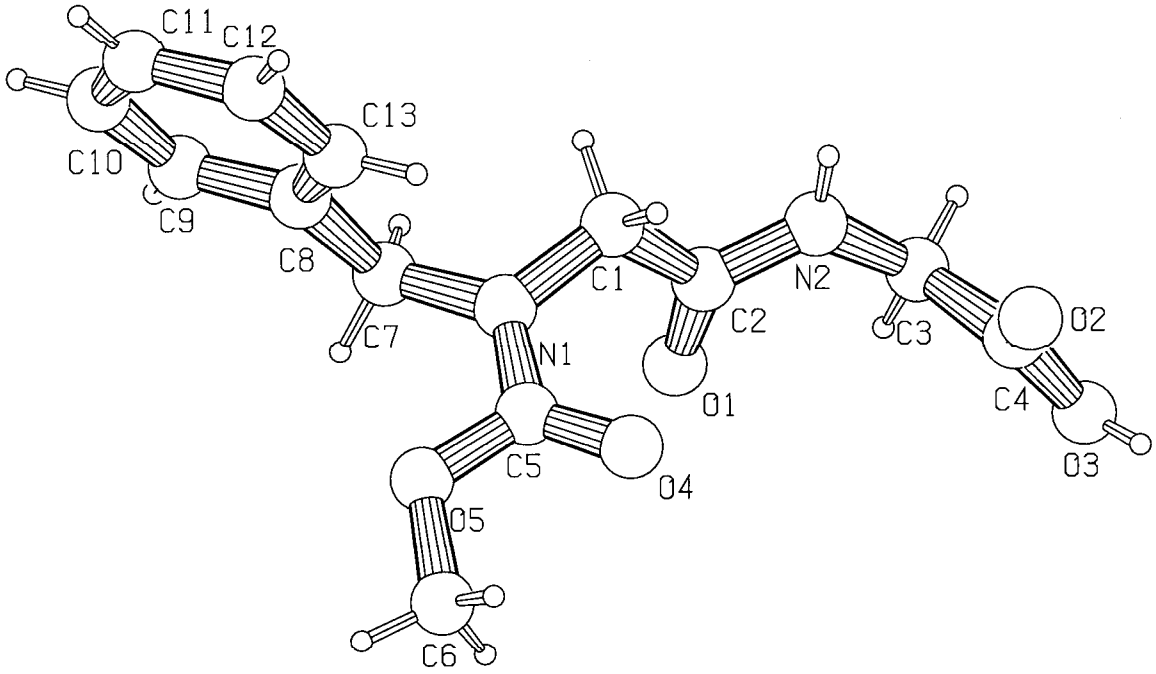


FIGURE 7. Crystal structure conformation (conventional *R*-factor = 0.1006) of an aspartame peptoid peptidomimetic.[2]

by a slight out-of-plane deformation of the coordination geometry at the amide nitrogen. The σ values (standard deviations, see Table 1) of the distributions for χ_C and χ_N in the proline amide bonds are larger than the corresponding values in the acyclic tertiary amides (Table 1), but in both fragments the σ value for χ_N is roughly three times the σ value for χ_C. The relatively larger flexibility for out-of-plane movements at the amide nitrogen compared with that at the carbonyl carbon atom is therefore nearly the same in both proline and acyclic tertiary amide bonds. Qualitatively similar results have been obtained by Dunitz and Winkler for the out-of-plane deformations of secondary amide bonds in crystal structures.[3]

In the solid state the distribution for the acyclic tertiary amide bond twisting parameter τ shows a maximum at 180˚ and secondary maxima at 177 and -177˚ (Fig. 5 (c)). It is unclear whether the secondary maxima observed here have a statistical significance. On the other hand, for the proline amide bonds a similar, but less pronounced distribution of τ is observed (Fig. 5 (f)). In support of these findings, τ values in the *in vacuo* optimized structures of Ac-NAla-NMe_2 and *N*,*N*-dimethylacetamide in C_1 symmetry correspond to either the highest or second-highest classes in the distribution in Fig. 5 (c) (see Table 3). The C5-N6 amide bond parameter τ in the α conformation of Ac-NAla-NMe_2 deviates even more from 180˚. The differences between the gas-phase out-of-plane parameter values (Table 3) and their preferential values in small-molecule crystal structures (Fig. 5) are probably related to the different aggregation states of the molecules. Analysis of the correlation between the out-of-plane parameters in the crystal structures may shed light on the detailed effects of the structured crystal structure environment on the geometry of the tertiary amide bonds, but is outside the scope of our work.

Recently, we determined the first crystal structure of a peptoid peptidomimetic, *viz*. a non-sweet analogue of the dipeptide artificial sweetener aspartame.[2] In this crystal structure the *trans* rotamer of the tertiary amide bond along the main chain is present (Fig. 7), meanwhile distinct rotamers resulting from rotation about the tertiary amide bond have been observed in the ^{1}H-NMR spectrum of the compound in DMSO.[2] The tertiary amide bond in the aspartame peptoid is part of a carboxylic acid methylester end-group of which the carbonyl oxygen atom O(4) accepts an intermolecular hydrogen bond from the secondary amide N(2)-H group of a neighbouring molecule (Fig. 7). The out-of-plane parameters for the tertiary amide group have the following values: χ_N = -4.6(7)˚ for N(1), χ_C = -1.0(7)˚ for C(5) and the twisting parameter τ = 174.0(4)˚. The estimated standard deviations are given in

parentheses. The values for the out-of-plane parameters χ_C and χ_N fall well within one standard deviation from the preferential value of zero for these parameters in the acyclic tertiary amide bonds in the solid state (Table 1). In addition, the τ value is just beyond one standard deviation away from 180˚. The out-of-plane geometry of the tertiary amide bond in the *peptoid* crystal structure thus is compatible with the geometrical behaviour of acyclic tertiary amide bonds in the solid state. This supports that acyclic tertiary amide bonds can be considered characteristic for the relatively rigid part of peptoids in the solid state.

Amide-Bond Geometries in Protein Binding Sites

The extreme values for the out-of-plane parameters χ_C, χ_N and τ in proline residues in protein binding sites (Table 3) correspond reasonably well with the extremes of the parameter values for proline amide bonds in the CSD (Table 1). In the sarcosine residues in the PDB the χ_N values are considerably lower than in the CSD. Given the limited variation in the sarcosine binding sites (only two clearly different sites) these values can not be considered representative for the behaviour of sarcosine residues in all binding sites. On the other hand, this limited variation in binding sites (such as for the cyclosporin A/cyclophilin complexes) might be exploited to study the 'flexibility' of specific binding sites and its influence on the (non)planarity of the amide bonds.

The out-of-plane deviations in a number of proline residues (χ_N = -45˚ in 1SCN:Pro449, resolution 1.9 Å)[16] and for some secondary amide carbon atoms (χ_C = -23 (1CDK:GlyI10) and χ_C = -31˚ (1CDK:GlyJ10), resolution 2.0 Å)[17] are considerably large, but are still within the estimated coordinate error of 0.3-0.4 Å, which is related to the resolution of the structures.[21-24] For instance, the proline nitrogen atom in 1SCN:Pro449 is 0.3 Å out-of-plane with the plane through its covalently bound atoms. These results, especially the χ_C values, are unexpected in view of the CSD results, but they cannot *a priori* be considered incorrect. It is, on the other hand, clear that depending on the protein and on the ligand, a non-planar geometry can be anticipated.

Influence of the Main Chain

With respect to the influence of the direction of the main chain in a peptoid on the non-planarity of the amide nitrogen coordination geometry, the protein-bound sarcosine residues provided a hint; the standard deviation in the main chain amide torsion angles, ω1

(mean value is 178.9, σ=1.7˚), is about half the standard deviation in torsion angles ω4, (mean -1.0, σ=3.6˚), that can be considered to describe the out-of-plane displacement of the side chain carbon atom with respect to the plane through C_α', C' and N (Fig. 4).[25] This suggests that at least in cyclosporin A complexed to a cyclophilin, the non-planarity of the tertiary amide nitrogen coordination geometry in a sarcosine residue can mainly be attributed to the position of the side chain methyl group.

As can be seen in Table 3, the ω4 torsion angles in the tertiary amide bonds in Ac-NAla-NMe_2 *in vacuo* differ more from zero than the ω1 torsion angle values differ from 180˚. The larger repulsion between the methylene and methyl groups (or between two methyl groups) in an eclipsed conformation, as compared with the repulsion between the less bulky carbonyl group and a methyl(ene) group in an eclipsed conformation, can be an explanation for this behaviour. In the solid state (and in the gas-phase[9]) this difference in repulsion can also be held responsible for a much larger C'NC(2) bond angle than the C'NC(1) bond angle both in proline (**2**) and acyclic tertiary amide bonds (**1**) (see Appendix B). The same effect will probably also have caused the slight lengthening of bonds (except C'-O') in the tertiary amide bonds with respect to the secondary amides as revealed by our results (Appendix B) and previous ones.[26] The crystallographically observed difference in the angles C'NC(1) and C'NC(2) was also reflected in molecular mechanics calculations using the CFF91 force field (data not shown). In conclusion, all results suggest that an asymmetric coordination geometry of the tertiary amide nitrogen can be ascribed to the repulsion of two sp3 hydridized carbon atoms in an eclipsed conformation along the amide bond.

Influence of the Environment

The differences in the preferred amide bond geometry in the gas phase and the solid state may be rationalized in view of the resonance model in Fig. 2. In small-molecule crystal structures hydrogen bond formation with, and protonation of the carbonyl oxygen atom in amide bonds has been shown to increase the resistance towards out-of-plane deformations of the trigonal amide nitrogen coordination geometry.[3] The carbonyl C'-O' bond length increases somewhat upon hydrogen bond formation and the C'-N' double bond character increases. The amide nitrogen atom rehybridizes towards an increased sp2 character (see the amide bond resonance forms in Fig. 2), with an associated reduction in the flexibility of its out-of-plane movement.[3,27] Also other groups/atoms approaching the carbonyl group, such as nearby nucleophiles,[4,5] can influence the electronic structure of the amide bond. These crystal field effects and also the different molecular frameworks in which the amide bonds

are embedded, can probably explain the preferred planar coordination geometry of tertiary amide nitrogen atoms in crystal structures (Fig. 4, Table 1). The out-of-plane deformation of the coordination geometry at the acyclic tertiary amide nitrogen was found to have a maximum at $\chi_N = 21^\circ$. In the gas phase a structured environment such as present in crystal structures and the intermolecular interactions associated with it, are absent. This allows for significant pyramidalization at the amide nitrogen atom (Table 3) to a χ_N of 30˚ even in local minimum energy conformations. This χ_N value is approximately half of the value encountered for a full pyramidal and tetrahedral sp3 coordination geometry. Our results as well as the use of the qualitative resonance model (Fig. 2) to explain them, are put in a different perspective by the gas electron-diffraction results by Mack and Oberhammer.[19] Their results stress the incapability of the resonance model to deal with the dynamics of a molecular system. Meanwhile, their and our conclusions are in good agreement: there is considerable flexibility at the tertiary amide nitrogen and a notable pyramidalization can occur at that site. For peptoids this implies that they intrinsically possess more possibilities to satisfy the requirements of the protein binding site than e.g. peptides. Given the sometimes strict requirements of a protein with respect to the stereochemistry (chirality) of ligands, this may be considered advantageous. For the amide bonds in the peptoid chains our results, in addition, suggest that in case of a *trans* or a *cis* rotamer, the first side chain carbon atom or the backbone C_α atom reveals the largest out-of-plane displacement, respectively.

EXPERIMENTAL SECTION

Substructure Searches in the Cambridge Structural Database

The geometry of the achiral amide moieties was determined in the asymmetric part of the conformational space characterized by the torsion angle value $0 \leq \omega 3(O'\text{-}C'\text{-}N\text{-}C(1)) \leq 180^\circ$. All other torsion angles were in the range -180 to 180˚. Conclusions can only be drawn for data in the *entire* conformational space; a different choice for its asymmetric part would result in a different distribution. Inversion of every fragment geometry results in the out-of-plane parameter values for the complementary asymmetric part (i.e. $-180 \leq \omega 3 \leq 0$). As can be seen in the corresponding definitions, the out-of-plane parameters then change sign (Fig. 4). The distribution for any of the three parameters in the *entire* conformational space is then the sum of the initial and its inverted distribution. In view of this, the left-hand side of the total distribution is the mirror image of the right-hand side, and as a result the average

value calculated for the total distribution always equals zero (or 180 in case of τ). Therefore, only the distributions itself show their modality.

The histograms contain 2N values, stemming from N independent fragments. There is not only an average value by definition (which results in the loss of a degree of freedom), but because of the symmetry of the distribution the average values (0 or 180) are also predetermined (loss of another degree of freedom). The total number of degrees of freedom therefore equals 2N (data points) minus N (relations between these points) minus 2 (because of the predetermined average value). The variation thus calculated for each distribution is roughly twice the variation for the same distribution with 2N independent values.

Geometry Optimizations of Ac-NAla-NMe_2 (4) using CFF91.

Starting models for **4** were built in INSIGHTII.[17] The backbone ϕ (C2-N3-C-C5) and ψ (N3-C-C5-N6) torsion angles (Fig. 6) were set at the corresponding MP2/6-31G* optimized values (See Table 3). The models were optimized to a final energy gradient of 0.001 kcal·mol^{-1}·$Å^{-1}$ using a dielectric constant of 1.

ACKNOWLEDGEMENT

Data for the amide bond geometries of glycine and proline residues in protein binding sites is courtesy of Drs Alfred Kostense. Critical discussions with Alfred Kostense and Dr Ed Moret are gratefully acknowledged. We thank the CAOS/CAMM Center, Nijmegen, the Netherlands for access to the Cambridge Structural Database System. *Ab initio* calculations were performed at the Janssen Research Foundation, Beerse, Belgium.

REFERENCES

1. Kruijtzer, J. A. W. and Liskamp, R. M. J., *Tetrahedron Lett.*, **1995**, *36*, 6969.
2. Boks, G. J., Boer, D. R., Kruijtzer, J. A. W., Liskamp, R. M. J., Tollenaere, J. P., Kroon, J. and Schouten, A., *unpublished results*, **1997**.
3. Dunitz, J. D. and Winkler, F. K., *Acta Cryst.*, **1975**, *B31*, 251.; Winkler, F. K. and Dunitz, J. D., *J. Mol. Biol.*, **1971**, *59*, 169.
4. Klebe, G., In: *Structure Correlation*, Vol. 2., eds. Bürgi, H.-B. and Dunitz, J. D., VCH, Weinheim 1994, p. 543-603.
5. Cieplak, A. S., In: *Structure Correlation*, Vol. 1., eds. Bürgi, H.-B. and Dunitz, J. D., VCH, Weinheim 1994, p. 205-302.

6. Pauling, L., In: *The Nature of the Chemical Bond*, 3d ed., Cornell University Press, Ithaca, NY, 1960, p. 281-282.
7. Allen, F. H. and Kennard, O., *Chem. Design Automation News*, **1993**, *8*, 31.; Allen, F. H., Davies, J. E., Galloy, J. J., Johnson, O., Kennard, O., Macrae, C. F., Mitchell, E. M., Mitchell, G. F., Smith, J. M., Watson, D. G. *J. Chem. Inf. Comput. Sci.*, **1991**, *31*, 187.
8. Bürgi, H.-B., *Acta Cryst.*, **1988**, B44, 445.
9. Moehle, K. and Hofmann, H.-J., *Biopolymers*, **1996**, *38*, 781.
10. Maple, J. R., Dinur, U., Hagler, A. T., *Proc. Natl. Acad. Sci. USA*, **1988**, *85*, 5350; Maple, J. R., Hwang, M.-J., Stockfish, T. P., Dinur, U., Waldman, M., Ewig, C. S., Hagler, A. T., *J. Comput. Chem.*, **1994**, *15*, 162.
11. IUPAC-IUB Commission on Biochemical Nomenclature, *J. Mol. Biol.*, **1970**, *52*, 1.
12. Dunitz J. D. and Bürgi, H.-B., In: *Structure Correlation*, Vol. 1., eds. Bürgi, H.-B. and Dunitz, J. D., VCH, Weinheim 1994, p. 23-68.
13. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Kennard, O., Shimanouchi, T. and Tasumi, M., *J. Mol. Biol.*, **1977**, *112*, 535.
14. Coll, M., Knof, S. H., Ohga, Y., Messerschmidt, A., Huber, R., Moellering, H., Ruessmann, L. and Schumacher, G., *J. Mol. Biol.*, **1990**, *214*, 597.
15. Steinmetz, A. C. U., Demuth, H.-U., Ringe, D., *Biochemistry*, **1994**, *33*, 10535.
16. Bossemeyer, D., Engh, R. A., Kinzel, V., Ponstingl, H., Huber, R., *EMBO J.*, **1993**, *12*, 849.
17. Biosym Technologies, 9685 Scranton Road, San Diego, CA 92121-2777 USA, 1995.
18. SPARTAN Version 4.0, Wavefunction Inc., 18401 Von Karman Ave., #370, Irvine, CA 92715 USA, 1995.
19. Mack, H.-G. and Oberhammer, H., *J. Am. Chem. Soc.*, **1997**, *119*, 3567.
20. Bludsky, O., Sponer J., Leszczynski, J., Spirko, V. and Hobza, P., *J. Chem. Phys.*, **1996**, *105*, 11042.
21. Chambers, J. L., Stroud, R. M., *Acta Cryst.*, **1979**, *B35*, 1861.
22. Dauber-Osguthorpe, P., Roberts, V. A., Osguthorpe, D. J., Wolff, J., Genest, M., Hagler, A. T., *Proteins*, **1988**, *4*, 31.
23. Wlodawer, A., Nachman, J., Gilliland, G. L., Gallagher, W., Woodward, C. J., *J. Mol. Biol.*, **1987**, *198*, 469.
24. Kossiokoff, A. A., Randal, M., Guenot, J., Eigenbrot, C., *Proteins*, **1992**, *14*, 65.
25. The torsion angles describing the conformation of the amide bond are related through the following equation: $(\omega1+\omega2) - (\omega3+\omega4) = 0 \bmod 2\pi$.[3] Substitution in the definition for χ_N in Fig. 4 will show that the alternative definition for $\chi_N=(\omega1-\omega4+\pi) \bmod 2\pi$ only differs in sign from the one previously defined.
26. Engh, R. A. and Huber, R., *Acta Cryst.*, **1991**, *A47*, 392.
26. Cieplak, A. S., *Struc. Chem.*, **1994**, *5*, 85.

Chapter 6

Putative Bioactive Conformations for a Fragment Characteristic of Peptoid Peptidomimetics

ABSTRACT

The conformational preferences of a fragment characteristic of peptoids (oligomers of *N*-substituted glycines), as encountered in small-molecule crystal structures and as part of ligands in protein complexes, were evaluated to assess energetically accessible backbone conformations of peptoids. We hypothesize that the conformational behaviour of peptoid residues in protein binding sites and in the solid state can be characterized by the conformational behaviour of the methylene bridge connecting two tertiary amide bonds. This methylene group acts as a ball-joint at which the amide bonds prefer a mutually perpendicular orientation. Inferences for side-chain positions and periodic (secondary) structures are discussed.

INTRODUCTION

Peptoids (oligomers of *N*-substituted glycines) form a class of peptidomimetics[1,2] that play an increasingly important role in the design of ligands targeted at therapeutically interesting proteins or other macromolecular receptors.[3,4] Biologically active peptoids have already been discovered for various systems[5,6] e.g. bovine pancreatic α-amylase and hepatitis A virus 3C protease.[5] Also, binding of a peptoid epitope to the transactivator-responsive element RNA (TAR) of the human immunodeficiency virus type 1 has been observed.[5] Using a combinatorial synthetic approach, and screening of a library of dimer and trimer peptoids carrying non-proteinogenic amino-acid side chains resulted in ligands for the α_1-adrenergic and μ-opiate receptor with affinities in the nanomolar range.[6] Structure-based design of peptoids for macromolecular receptors has so far been restricted to exploiting the topological similarity of peptoids and the peptides they intend to mimic (see Figure 1, Chapter 1). Experimental structural chemical knowledge on the structure and conformational behaviour of peptoids is still very limited. Recently, the first crystal structure of a peptoid *viz.* an analogue of the dipeptide artificial sweetener aspartame, has been reported.[7] In addition, *cis/trans* isomerization of the tertiary amide bonds in peptoids has been observed in NMR experiments in various solvents at ambient temperature.[7,8] Also the tertiary amide nitrogen atom was found to be the site at which considerable pyramidalization as well as a planar coordination geometry is possible, depending on the environment (see Chapter 5). Here we focus on preferential conformations in small-molecule crystal structures of a fragment that is believed to be characteristic for peptoids, and on *N*-methylglycine (sarcosine) residue conformations in protein binding sites. Conformational preferences as reflected in preferred torsion angles in selected (open-chain) fragments in small-molecule crystal structures have been found to correspond to torsion angle preferences of ligands in protein complexes.[9,10] The influence of the structured protein environment on the conformations of ligands can, according to these findings, thus be mimicked by the structured environment in small-molecule crystal structures. By applying this line of thinking to the conformational preferences of characteristic peptoid fragments in the solid state we might be able to propose putative bioactive conformations for peptoids.

Peptoids are built up from amino-acid based *N*-substituted glycines connected *via* tertiary amide bonds (Fig. 1). Glycine and proline are peptoid residues in themselves, this implies that in addition to tertiary amide bonds also secondary amide bonds can (i.e. in the case of glycine residues in the chain) be present. The atom labelling scheme adopted for

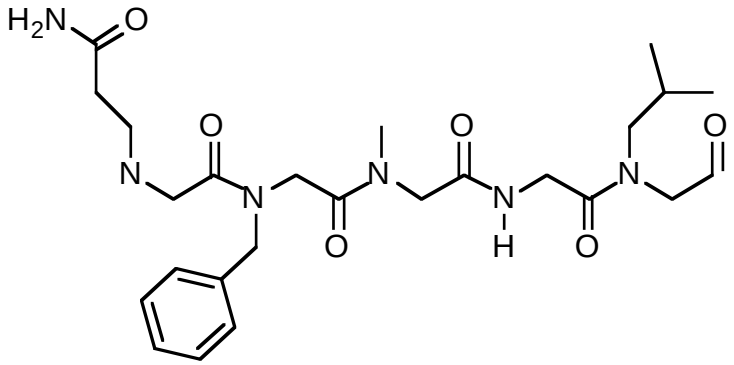
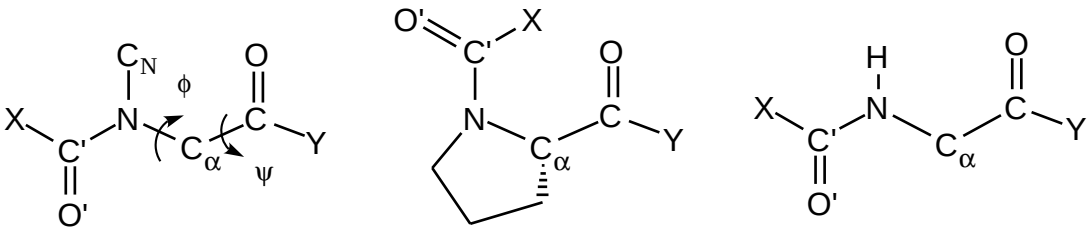


FIGURE 1. A general peptoid sequence (-NGln-NPhe-NAla-Gly-NLeu-) is indicated. The prefix 'N' in the peptoid monomer names denotes that the side chain has been transferred to the amide nitrogen of the corresponding amino acid (see Appendix A).

peptoids is, as far as the peptoid backbone is concerned, identical to the nomenclature for peptides (see Appendix A).[11] As a result, the definitions of the backbone torsion angles ϕ, ψ and ω are the same for peptides and peptoids. The side chain carbon atom attached to the amide nitrogen atom is denoted C_N, when no further specification is required (Appendix A). The Cambridge Structural Database[12] (CSD, version 5.12, 160,091 entries) and the Brookhaven Protein Data Bank[13] (PDB, November 1996) were searched for a molecular substructure characteristic of *N*-substituted glycines.



1 **2** **3**

FIGURE 2. Peptoid residues: (**1**) a generic peptoid residue substructure, a proline (**2**) and glycine (**3**) residue

In the PDB, the backbone conformation of sarcosine (*N*-methylglycine) residues that are part of ligands in protein-ligand complexes was analysed. In substructure searches in the CSD substructure **1** (Fig. 2) was considered to be representative for the flexible part of peptoid

oligomers, with the exception of proline (**2**) and glycine (**3**) residues (Fig. 2). The terminal amide-resembling groups need not be identical to a tertiary (or secondary) amide group as encountered in peptoids (see Fig. 1). In this respect, we want to exclude fragments with a tetrahedrally coordinated amide nitrogen atom in the N-terminal amide bond (e.g. as may be the case for a urea group, see Chapter 5) and to include C-terminal carboxylic acids, ethers and primary carboxamide groups -when present- into our searches. Fragments with a considerable out-of-plane deformation of the amide moiety (more than 3σ away from the average crystallographic values for the out-of-plane parameters[14] χ_C, χ_N and τ, see Chapter 5) were therefore individually evaluated and discarded whenever the fragments were considered to be inappropriate to represent a peptoid substructure. As a consequence we identified energetically accessible conformations of fragments, that are considered characteristic for peptoid residues in a highly structured molecular environment. In addition, we propose conformational features that might be characteristic for the bioactive conformations of peptoids.

METHODS AND MATERIALS

Substructure Searches in the Cambridge Structural Database

Searches were performed using QUEST 2.17 and the results were analysed with VISTA 2.3, both of which are part of the Cambridge Structural Database System.[12] C_α in substructure **1** (Fig. 2) is requested to be a secondary carbon atom, C_N is any 'acyclic' carbon atom. X and Y denote any non-metal or non-hydrogen atom. No other covalent bonds were allowed between atoms in the fragment than the ones indicated in the structural diagrams (Fig. 2). Hydrogen bonds accepted by the carbonyl oxygen atoms were permitted, but covalent coordination with e.g. metal atoms was excluded. In cases where the fragment is required to be a part of a cyclic molecule, the carbonyl carbon atoms on either end were requested to be of a 'cyclic' nature. Structures with a conventional R-factor ≤ 0.10 and a covalent bond tolerance of 0.4 Å (in establishing connectivity) were included in the search. Structures that were error-flagged, contained disorder or were part of polymers were excluded. Duplicate structures were also discarded. Fragments that were outliers in a bond length or bond angle value by more than four times the standard deviation from the mean value of the corresponding parameter in the set were discarded. The conformation of the achiral fragment was determined in an asymmetric part of the fragment's conformational space.[15]

Sarcosine (*N*-Methylglycine) as Ligand Residue

Crystallographically determined protein-ligand complexes containing sarcosine (*N*-methylglycine) as part of the bound ligand were retrieved from a survey of the Brookhaven Protein Data Bank[13] (PDB, November 1996). The search was restricted to well resolved structures with a high resolution (≤ 2.5 Å) and an R-factor ≤ 0.20. In addition, all relevant atom positions (as given in substructure **1** in Fig. 2) should have been explicitly determined. In contrast to the substructure searches in the CSD, where the fragment conformation was determined in an asymmetric part of the conformational space, here residue conformations were evaluated as actually present in the protein binding sites.

RESULTS

Small-Molecule Crystal Structures

From the survey of the Cambridge Structural Database (CSD) we collected 14 crystal structures that contained substructure **1** (see Fig. 2) as part of an acyclic molecule. Two fragments with "outlier" τ values (see Chapter 5) of -17 and -149°, which were associated with extensive overlap of a second crystallographically unique fragment in the same structures, were discarded. As a result 17 unique peptoid fragments were located in these structures.[16] The C-terminal groups of the fragments were carboxylic acids, esters, secondary or tertiary amide groups. X was a carbon or oxygen atom (Fig. 2, **1**). The conformation of the ϕ (C'-N-C_α-C) and ψ (N-C_α-C-Y) torsion angles in each fragment is depicted in the asymmetric part (-180≤ϕ≤0°, -180<ψ≤180°) of a Ramachandran-like plot (Fig. 3 (a)).[17] Inversion of every point through the origin yields the mirror image conformation. Two clusters of conformations are present, which can be characterized by the central (ϕ,ψ) values (and their outer limits) -85±23, 175±21° and -88±6, 2±16°. Representative conformations are depicted in Figs 4 (a) and (b). The first cluster will be referred to as the ψ-*trans* cluster, the latter as the ψ-*cis* cluster. Only two fragments in the set contain a tertiary amide group at the C-terminus; their conformations are part of the ψ-*trans* cluster. The ψ-*cis* cluster contains one fragment with a C-terminal carboxylate group and four fragments with a C-terminal secondary amide group. In the latter four fragments the conformation is stabilized by an intramolecular hydrogen bond involving the secondary amide nitrogen. In the two clusters both *cis* and *trans* amide bonds, when viewed along the main chain of the molecules, are present.

We located 61 unique peptoid monomer fragments (**1**) in cyclic molecules in the CSD with ring sizes varying from seven to 33 atoms in a ring.[18] Two fragments with χ_N values (26 and 27˚, respectively) slightly beyond the 3σ criterion $|\chi_N|>25.5$˚ (See Chapter 5), were retained in the set. The only seven-membered ring present was a imidazodiazepine analogue.[19] All other rings were peptide analogues. In the seven and nine-membered rings only *cis* amide bonds when seen along the main chain are present in the fragments. For ring sizes of 12 or larger, both *cis* and *trans* amide bonds appear in the fragments. The conformations of the fragments with ring sizes varying from 12 to 33 atoms are depicted as open circles in Fig. 3 (b). Conformations of the fragments in the seven- and nine-membered rings are shown as a cross and as filled circles in the same Figure, respectively. The conformations are depicted in the asymmetric part of the Ramachandran plot. Three clusters of conformations are present in the 12- to 33-membered rings. (i) A cluster with the central (ϕ,ψ) values -97±29, 177±19˚, which is called the ψ-*trans* cluster, because of its resemblance to the cluster with similar ϕ,ψ values for the acyclic peptoid fragments. (ii) The cluster characterized by the (ϕ,ψ) values -69±6, 141±8˚, is denoted the ϕ-(-)*gauche* cluster. (iii) A third cluster is present within the (ϕ,ψ) values -128±7, 67±3˚, which is called the ψ-(+)*gauche* cluster. Characteristic conformations for the latter two clusters are depicted in Figs 4 (c) and (d). The seven-membered imidazodiazepine ring contains a peptoid fragment with (ϕ,ψ) torsion angle values -78, 71˚. The fragment conformations in the nine-membered rings were either nearly eclipsed along C_α-C with (ϕ,ψ) values -74, 9˚ and -58, -7˚, just outside the previously identified ψ-*cis* cluster in the acyclic compounds, or roughly follow the imaginary line between the ψ-(+)*gauche* and ϕ-(-)*gauche* clusters of the larger rings (see Fig. 3 (b)).

Protein Binding Sites

In the Brookhaven Protein Data Bank (PDB) ten protein complexes containing sarcosine as part of the bound ligand were identified (see Table 1). Note that the variation in the protein binding sites of the sarcosine residues is rather limited; eight out of the ten complexes are cyclophilins complexed with cyclosporin A or modifications thereof.[21-24,26-28] In addition, a structure of cyclosporin A in complex with the antigen binding Fab part of a murine IgG_1-κ antibody[25] and a complex of creatine amidinohydrolase (creatinase) complexed with carbamoyl sarcosine were present.[20] The backbone conformations of the 29 sarcosine residues in the various cyclophilin-cyclosporin complexes cluster around (ϕ,ψ) 125±7, -75±9˚ (see Fig. 5). The sarcosine residue in cyclosporin A complexed with the

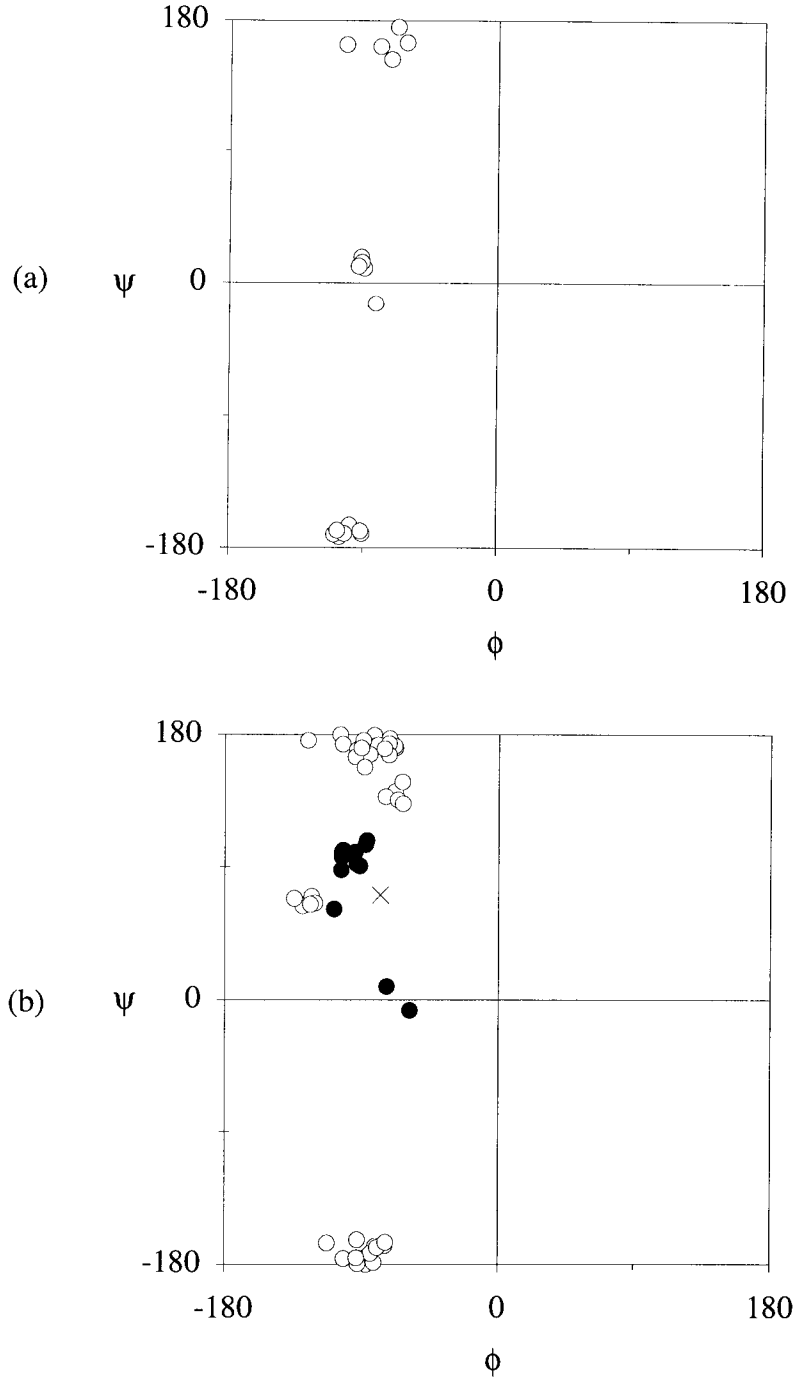


FIGURE 3. Conformer distribution for substructure **1** as present in the Cambridge Structural Database (CSD).[12] Indicated are the ϕ (C'-N-C_α-C) and ψ (N-C_α-C-Y) torsion angles in a Ramachandran-type plot.[17] (a) 17 Acyclic fragments. (b) 61 Cyclic fragments; ring sizes from 12 to 33 atoms are given as open circles. The seven and nine-membered rings are given as a cross and as filled circles, respectively.

antibody Fab fragment, on the other hand, was found at (ϕ,ψ) values 84, -173˚ in the Ramachandran plot (Fig. 5). The two carbamoyl sarcosines bound to creatine amidinohydrolase with (ϕ,ψ) values 140, 4˚ or 140, 179˚ for one residue and 149, 12˚ or 149, -167˚ for the other residue. The ambiguity in determining the ψ torsion angle exists, because the C-terminal group is a carboxylate group, which, in addition, is non-planar Only the (ϕ,ψ) values with ψ values close to a ψ-*cis* conformation are shown in Fig. 5. The N-terminal amide group of these two sarcosine residues is part of a carbamoyl group with a tetrahedral coordination geometry at the tertiary amide nitrogen atom with χ_N values of 60 and 64˚. The trigonal coordination of the carbonyl carbon atom in one of the sarcosines, was found to be χ_C=-42˚ out-of-plane (see Chapter 5).

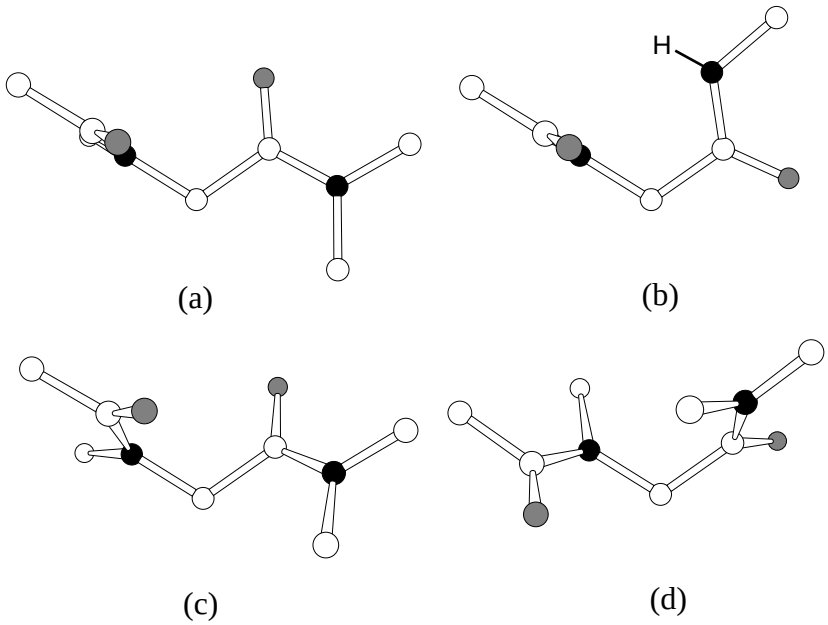


FIGURE 4. Experimentally determined conformations for substructure **1** as characterized by the clusters in Fig. 3. (a) ψ-*trans*, (ϕ,ψ) is -85, 175˚, (b) ψ-*cis*, (ϕ,ψ) is -88, 2˚, (c) ϕ-(-) *gauche*, (ϕ,ψ) is -69, 141˚, (d) ψ-(+)*gauche* , (ϕ,ψ) is -128, 67˚. Nitrogen atoms are shown in black, oxygen atoms in grey and carbon atoms in white. The fragments are shown in such a way that the N, C_α and C atoms are in the plane of the paper.

TABLE 1. Crystallographically determined protein-ligand complexes containing a sarcosine (*N*-methylglycine) residue as part of the bound ligand

Protein	Ligand	No. of sarcosines	PDB entry	Resolution (Å)	R-factor	Ref.
Creatine amidinohydrolase (*Pseudomonas Putida*)	Carbamoyl sarcosine	2	1CHM	1.9	0.177	20
Human cyclophilin A	Cyclosporin A	1	1CWA	2.1	0.167	21
Human cyclophilin A	Cyclosporin analogue	1	1CWB	2.2	0.162	22
Human cyclophilin A	Cyclosporin analogue	1	1CWC	1.86	0.177	23
Human cyclophilin B	Cyclosporin analogue	1	1CYN	1.85	0.160	24
Murine IgG1-κ antibody Fab fragment	Cyclosporin A	1	1IKF	2.5	0.164	25
Human cyclophilin A	Cyclosporin analogue	1	1MIK	1.76	0.175	26
Human cyclophilin A	Cyclosporin A	10	2RMA	2.1	0.170	27
Human cyclophilin A	Cyclosporin analogue	10	2RMB	2.1	0.184	27
Murine cyclophilin C	Cyclosporin A	4	2RMC	1.64	0.197	28

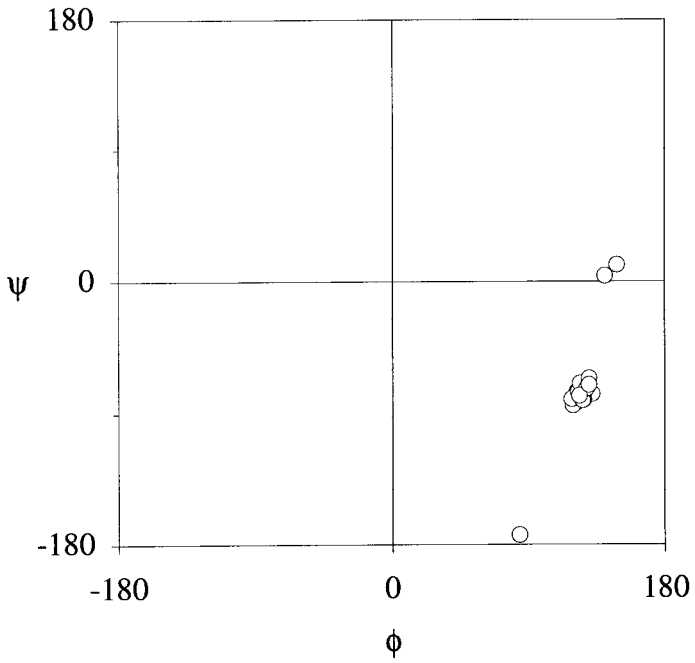


FIGURE 5. Conformer distribution for sarcosine (*N*-methylglycine) residues as present in binding sites in crystallographically determined protein complexes in the Brookhaven Protein Databank (PDB).[13] Indicated are the ϕ (C'-N-C_α-C) and ψ (N-C_α-C-Y) torsion angles.

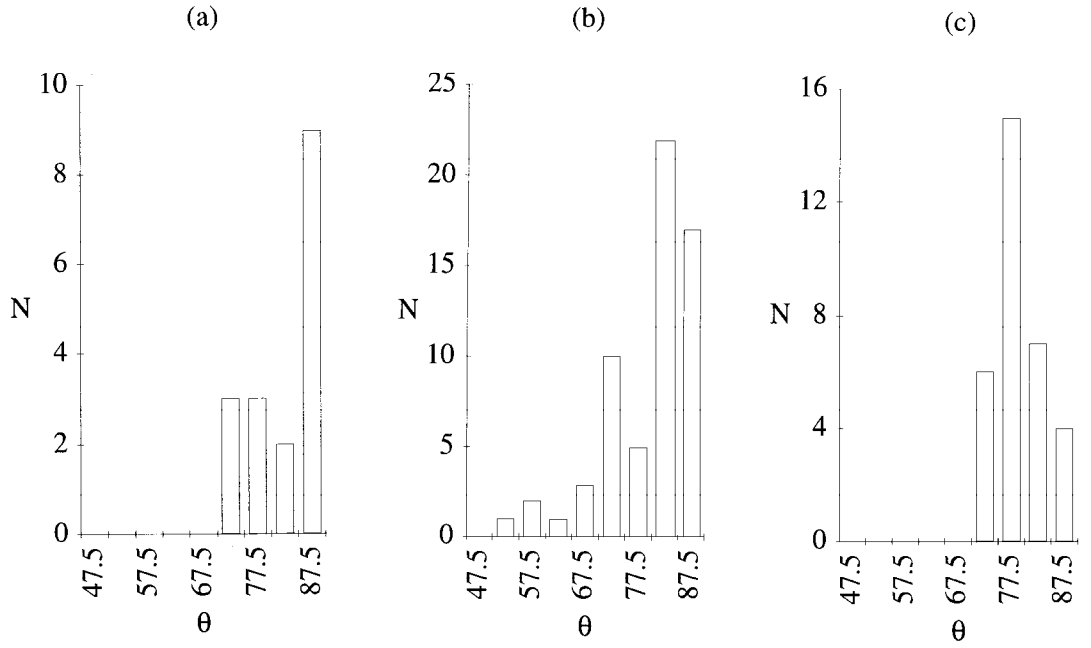
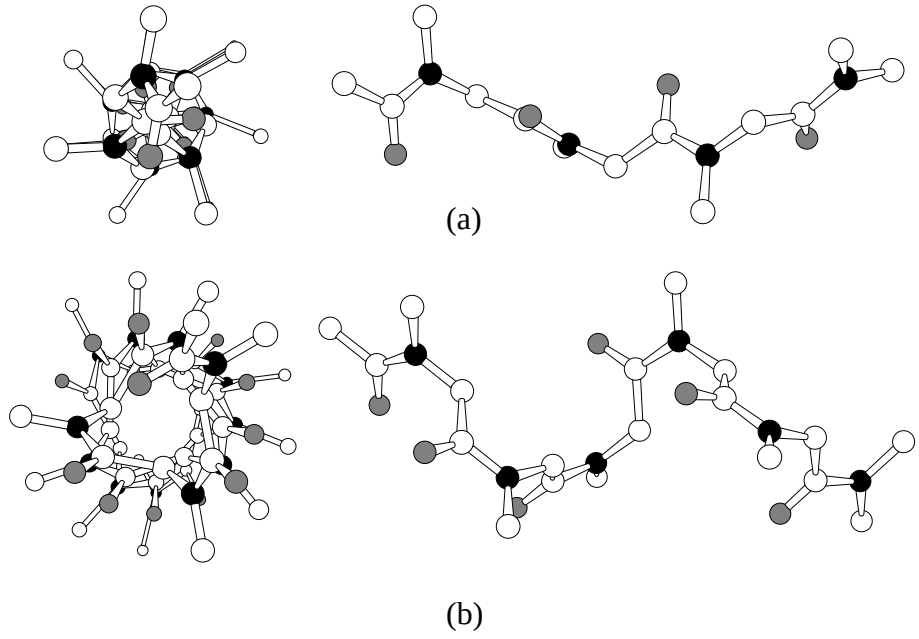


FIGURE 6. Distributions for the interplane angle θ (0≤θ≤90˚)between the least-squares planes through two amide moieties adjacent to C_α. Substructure **1** in (a) acyclic and (b) cyclic molecules in the CSD[12] and (c) sarcosine residues in the PDB.[13] Class widths are 5˚.

The Mutual Orientation of Amide Bonds

Acyclic tertiary amide bonds in small-molecule crystal structures were found to be planar on average (Chapter 5). Therefore, the characteristic peptoid fragment as exemplified by substructure **1** might be approximated by two ideal planes connected by a methylene bridge. The backbone conformations of the peptoid fragments can in view of this also be described by the angle θ (0 ≤ θ ≤ 90˚) which is the interplane angle between the least-squares planes through the atoms in plane with the amide group (X, C', O', N, C_α, C_N) and the carbonyl group (C_α, C, O, Y) adjacent to C_α (Fig. 2). The distribution of the interplane angle θ in the acyclic and cyclic fragments and in the protein bound sarcosine residues is given in Fig. 6. The interplane angle θ in the acyclic fragments ranges from 71 to 90˚, with a pronounced preference for the region 85-90˚ (Fig. 6(a)). *Cis* and *trans* amide bonds (at the N-terminal side of the fragment) are present over the entire range. In the cyclic fragments (Fig. 6(b)) twice as much *cis* amide bonds as the number of *trans* amide bonds are present. The interplane angle θ varies over a larger range: 52 to 90˚, with a preference for the range 80-90˚ for both *cis* and *trans* amide bonds. The lower interplane angle interval in the cyclic

fragments from 52 to 78˚ is associated with *cis* amide bonds only (including the seven and nine-membered rings), and has a maximum at 70-75˚. In the 32 sarcosine residues in protein binding sites the interplane angle θ ranges over the interval 72 to 90˚, with preferential values of 75-80˚ (Fig. 6 (c)). These values are, similar to the populations in Fig. 5, strongly biased by the many cyclophilin-cyclosporin complexes. The interplane angle values in the carbamoyl sarcosines complexed with creatine amidinohydrolase[20] (*vide infra*) are θ=85 and 72˚. The latter value corresponds to the residue with an extreme out-of-plane coordination geometry of the amide carbonyl carbon atom (χ_C=-42˚). In all sarcosine residues only *trans* amide bonds were present at the N-terminal side of the fragments.



(a)

(b)

FIGURE 7. Helical wheel plots and (partial) side views of an undecamer consisting of *N*-methylglycine residues with a ψ-*trans* conformation (ϕ,ψ) -85, 175˚ with (a) *trans* amide bonds in a left-handed helix and (b) *cis* amide bonds in a right-handed helix. Subsequent side chains are rotated by ca. -110 (a) or +110˚ (b) along the helix axis with this specific backbone conformation.

Modelled Periodic Structures

To illustrate the relative side-chain directions in a peptoid oligomer constructed from energetically accessible peptoid monomer conformations, helical peptoid undecamers were built with a ψ-*trans* conformation (Fig. 7). Helices with either *trans* or *cis* amide bonds could

be modelled without severe sterical hindrance. When viewed along the helical axes (see Fig. 7), the side chains in subsequent residues are rotated by -90 to -120 or +90 to +120˚, depending on the exact backbone conformation and the handedness of the helix. In the helix with only *trans* amide bonds (Fig. 7(a)) the side chains, the carbonyl groups as well as the C_α carbon atoms seem to be exposed. The N-C_N bonds are oriented more or less perpendicular to the helical axis. In the helix with only *cis* amide bonds the C_α atoms are buried, the side chains and the neighbouring carbonyl groups are exposed and are held closely together by a *cis* amide bond. Subsequent full geometry optimization of these conformations using the CFF91 force field,[29] did not severely distort the helices; the backbone torsion angles changed from (ϕ,ψ) -85, 175˚ to -81, 167˚ and -75, 167˚ for the helices with *trans* and *cis* amide bonds, respectively. The dipole moments in these helices coincide with the helical axes and are directed from the C- towards the N-terminus. When using the CFF91 partial atomic charges, the most compact helix, with *cis* amide bonds, possesses a dipole moment of 38.7 D, whereas the dipole moment of the helix with *trans* amide bonds amounts to 10.5 D.

DISCUSSION

All clusters of the solid state and bioactive conformations of peptoid fragments can be characterized by the interplane angle θ of the least-squares planes through adjacent amide moieties, which reveal a (strong) tendency for a perpendicular orientation of two adjacent amide moieties (Figs 4 and 6), even when the amide nitrogen is the site of a considerable pyramidalization (e.g. in the case of the creatine amidinohydrolase complex 1CHM). This behaviour seems to be characteristic of energetically accessible peptoid monomer conformations, also in their bioactive state. In the solid state ('acyclic' and 'cyclic') and in protein binding sites, θ ranges preferentially within a 70 to 90˚ interval for both *cis* and *trans* amide bonds along the main chain. The degree of distortion from a perpendicular orientation of two adjacent amide bonds in peptoids will probably depend on the peptoid sequence and the requirements of the protein binding site. Also, cyclization of peptoids will probably induce such as distortion, by analogy to the cyclic molecules in the solid state (Fig. 6(b)).

The so-called ψ-*trans* backbone conformation is found as the bioactive conformation of a sarcosine residue in cyclosporin A bound to a murine antibody Fab fragment, as well as in the solid state conformations of the flexible peptoid substructure (**1**) in cyclic and acyclic molecules. This recurring peptoid backbone conformation (*vide infra*) is characterized by a perpendicular orientation of the N-terminal amide moiety ($\phi \approx \pm 90$˚) and the plane through N-

C_α-C, and a coplanar orientation of the C-terminal amide moiety ($\psi \approx 180^\circ$) to the same plane (Fig. 4(a)). Peptoid fragments might mimic in this respect the conformational behaviour of a diphenylmethyl fragment in the solid state.[15] The phenyl rings that are connected by a methylene bridge also prefer a mutually perpendicular orientation; one phenyl ring perpendicular to the plane through the connecting single bonds and one phenyl ring parallel to that. A 180˚ rotation of one of these phenyl rings around the benzyl single bond leaves its orientation unaltered. In peptoids, a 180˚ rotation of ϕ or ψ, starting from a ψ-*trans* conformation (Fig. 4 (a)), results in another ψ-*trans* or ψ-*cis* conformation, respectively.

A ψ-*cis* conformation in acyclic molecules was only encountered when the nitrogen atom of the C-terminal amide moiety was involved in an intramolecular hydrogen bond or when the C-terminus was a carboxylic acid group. Steric hindrance between side chain atoms of the succeeding residue in the chain and atoms from the N-terminal amide moiety (Fig. 4 (b)) prevents the C-terminal tertiary amide bonds to adopt this orientation. A ψ-*cis* conformation stabilized by an internal hydrogen bond may, however, be energetically accessible in case of a C-terminal secondary amide bond (from a succeeding glycine residue). For a planar carboxylate group at the C-terminus the ψ-*trans* and ψ-*cis* conformations are identical. The carboxylic acid methyl ester in the crystal structure of the aspartame peptoid analogue adopts a ψ-*trans* conformation (Fig. 7 in Chapter 5).[7] The two fragments in the nine-membered cyclic molecules which also adopt a ψ-*cis* conformation (Fig. 3 (b)), contain proline amide bonds at their C-terminal side.

The ψ-(+)*gauche* and ϕ-(-)*gauche* conformations seem to form the endpoints of a conformational transition between two staggered conformations, certainly with respect to the interplane angle θ. The fragment conformations in the nine-membered rings that lie in between these staggered conformations (Fig. 3 (b)) may then be considered as intermediate conformations to a perpendicular orientation of adjacent amide bonds along the 'reaction coordinate' θ. The ϕ-(-)*gauche* conformation is, in addition, very close to the ψ-*trans* conformation, see Figs 4 (c) and (a), respectively. Because of the symmetry of the conformational space, an equivalent situation applies for the lower-right quadrant of the Ramachandran plot.

In addition to the peptoid monomer substructure **1** depicted in Fig. 2, a peptoid chain can also contain glycine and proline residues. A glycine residue could, besides its possibility to form intramolecular hydrogen bonds, facilitate a ψ-*cis* conformation in a preceding residue in the peptoid chain (Fig. 4 (b)). The lack of a side chain at the glycine nitrogen atom will

probably also result in more energetically accessible conformations compared to the generic peptoid (*N*-substituted glycine) residue. As a peptoid backbone is (except for proline residues) devoid of chiral centres, the energetically accessible peptoid monomer conformations occur over a wider range in the Ramachandran plot than in case of α-amino acids.[17] The positions that are occupied by the peptoid conformations in the left-hand side of the Ramachandran plot correspond to low energy peptide monomer conformations. The centrosymmetry of the peptoid Ramachandran plot is, however, absent in peptides. Perhaps even more than in peptides, the introduction of L or D-proline residues might have an important role in restricting the peptoid backbone to a specific region of the available conformational space. Proline residues in the solid state conformations of acyclic molecules also prefer a nearly perpendicular orientation of two adjacent amide bonds with a preference of θ for 80-85˚ (unpublished results) and thereby reflect the conformational behaviour as seen for generic peptoid monomers. The work described here is currently being extended to include proline and glycine residues in the CSD and in protein-peptide complexes as well.

Recently, Moehle and Hofmann[30] published local minimum energy conformations of the model peptoid Ac-NAla-NMe_2 at various *ab initio* quantum chemical and (semi)empirical levels *in vacuo*. This was followed by molecular dynamics studies in solution and *in vacuo*.[31] They localized a total of three conformations (six after including the symmetry of the conformational space), which were local minimum conformations *in vacuo* either with *cis* or *trans* amide bonds, using both *ab initio* quantum chemical (including electron correlation) as well as (semi)empirical calculations. The conformations were denoted $C_{7\beta}$ (ϕ,ψ) -128.2, 77.0˚, α_D (ϕ,ψ) 74.2, -175.6˚ and α (ϕ,ψ) -54.7, -47.2˚.[32] The interplane angle values were θ=74, 87 and 79˚, respectively. The first two conformations correspond remarkably well with the ψ-(+)*gauche* ((ϕ,ψ) -128±7, 67±3˚) and ψ-*trans* ((ϕ,ψ) -85±23, 175±21˚ and -97±29, 177±19˚) conformations of peptoid fragments encountered in the solid state. Such a correspondence between preferential torsion angle values in crystal structures and local minimum energy conformations for corresponding model compounds in high quality quantum chemical calculations *in vacuo* has been observed before.[33] The α conformation was, however, not encountered in our results. And on the other hand, the ψ-*cis* and ϕ-(-)*gauche* conformations did not correspond to the *in vacuo* local minimum conformations. In molecular dynamics simulations *in vacuo* at 300 K using the CHARMm22 force field the $C_{7\beta}$ and α_D conformations were found to be stable during the simulation (200 ps).[31] The region corresponding to the (shortest) transition path between the two conformations was also significantly occupied when a *trans* amide bond was present. In our

results this region was occupied by the ϕ-(-)*gauche* cluster and many of the peptoid fragment conformations in the nine-membered rings. In molecular dynamics simulations in solution (300 K) only the α_D conformation remained for both *cis* and *trans* amide bonds.

In summary, all bioactive and solid state conformations encountered for peptoid fragments as well as the gas phase conformations of a model peptoid[30] can be characterized by the interplane angle θ, that reveals a pronounced tendency for a perpendicular orientation of two adjacent amide moieties. Due to the absence of hydrogen bond donor groups in the peptoid backbone the formation of stable conformations that are held together by multiple hydrogen bonds, such as in peptides, can probably be excluded. Only in the vicinity of glycine residues locally hydrogen bond stabilized conformations can be expected for peptoids. The nearly perpendicular orientation of the two adjacent amide bonds might be as characteristic for peptoid oligomer conformations in the bioactive and the solid state, as secondary structure motifs are in peptides. The experimentally determined peptoid backbone conformations thereby appear to be mainly restricted to the upper-left and lower-right quadrants of the Ramachandran plot.

The relative side chain orientations in neighbouring peptoid residues in a chain, while adopting the experimentally observed peptoid fragment conformations, cannot be readily assessed. It depends e.g. on the *trans* or *cis* conformation of the connecting amide bond, the degree of pyramidalization at the amide nitrogen, as well as the combination of specific conformations in the residues. Even then, only the direction of the N-C_N bond is thereby determined, which does not completely describe the conformation of the side chain. The ψ-*trans* conformation, when employed to build an undecamer peptoid, remained intact upon geometry optimization *in vacuo*. It can, however, not be stated that any of the experimentally encountered peptoid fragment conformations can result in energetically stable periodically repeated structures in a condensed environment, especially not in the absence of internal hydrogen bonds or other stabilizing interactions. On the other hand, the modelled helices prompt a number of intriguing suggestions. It might, for instance, be speculated that alkylating the C_α carbon atom in peptoids opposes the formation of *cis* amide bonds. The remarkable difference in dipole moments of the modelled helices probably reflects the influence of a *cis* or *trans* amide bond on the dipole moment also of smaller peptoids. Given the general importance of dipole moments in intermolecular interactions, the facile *cis/trans* isomerization of amide bonds in peptoids, might have consequences for the recognition of

peptoids by macromolecular targets. In this respect, alkylating the C_α carbon atom might have a dual role. A more elaborate approach than the one employed here to study these effects would be to combine the experimental results presented here with theoretical conformational analysis of peptoid oligomers and thereby perhaps be helpful in selecting the most likely candidates for bioactive peptoid conformations.

ACKNOWLEDGEMENT

The data for the proline and glycine residues constituting ligands in protein binding sites were kindly provided by Drs Alfred Kostense. We thank Alfred Kostense and Dr Ed Moret for critically reading the manuscript. We acknowledge the CAOS/CAMM Center, Nijmegen, the Netherlands for access to the Cambridge Structural Database System.

REFERENCES

1. Giannis, A. and Kolter, T., *Angew. Chem. Int. Ed. Engl.*, **1993**, *32*, 1244.
2. Moore, G. J., *Trends Pharmaceutical Sci.*, **1994**, *15*, 124.
3. Kessler, H., *Angew. Chem. Int. Ed. Engl.*, **1993**, *32*, 543.
4. Liskamp, R. M. J., *Angew. Chem. Int. Ed. Engl.*, **1994**, *33*, 633.
5. Simon, R. J., Kania, R. S., Zuckermann, R. N., Huebner, V. D., Jewell, D. A., Banville, S., Ng, S., Wang, L., Rosenberg, S., Marlowe, C. K., Spellmeyer, D. C., Tan, R., Frenkel, A. D., Santi, D. V., Cohen, F. E. and Bartlett, P. A., *Proc. Natl. Acad. Sci. USA*, **1992**, *89*, 9367.
6. Zuckermann, R. N., Martin, E. J., Spellmeyer, D. C., Stauber, G. B., Shoemaker, K. R., Kerr, J. M., Figliozzi, G. M., Goff, D. A., Siani, M. A., Simon, R. J., Banville, S. C., Brown, E.G., Wang, L., Richter, L. S. and Moos, W. H., *J. Med. Chem.*, **1994**, *37*, 2678.
7. Boks, G. J., Boer, D. R., Kruijtzer, J. A. W., Liskamp, R. M. J., Tollenaere, J. P., Kroon, J. and Schouten, A., *unpublished results*, **1997**.
8. Kruijtzer, J. A. W. and Liskamp, R. M. J., *Tetrahedron Lett.*, **1995**, *36*, 6969.
9. Klebe, G. and Mietzner, T., *J. Comput.-Aided Mol. Design*, **1994**, *8*, 583.
10. Klebe, G., *Persp. Drug Discov. Design*, **1995**, *3*, 85.
11. IUPAC-IUB Commission on Biochemical Nomenclature, *J. Mol. Biol.*, **1970**, *52*, 1.
12. Allen, F. H. and Kennard, O., *Chem. Design. Automation News*, **1993**, *8*, 31.; Allen, F. H., Davies, J. E., Galloy, J. J., Johnson, O., Kennard, O., Macrae, C. F., Mitchell, E. M., Mitchell, G. F., Smith, J. M., Watson, D. G., *J. Chem. Inf. Comput. Sci.*, **1991**, *31*, 187.
13. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Kennard, O., Shimanouchi, T. and Tasumi, M., *J. Mol. Biol.*, **1977**, *112*, 535.
14 (a). Dunitz, J. D. and Winkler, F. K., *Acta Cryst.*, **1975**, *B31*, 251.; (b) Winkler, F. K. and Dunitz, J. D., *J. Mol. Biol.*, **1971**, *59*, 169.

15. Dunitz, J. D. and Bürgi, H.-B., In: *Structure Correlation*, Vol. 1., eds. Bürgi, H.-B. and Dunitz, J. D., VCH, Weinheim 1994, p. 23-68.
16. CSD refcodes: BCPSBZ, BPROSA, BPROSA01, BXSASA, JOWVUM, PAGTIA, PAGTOG, PAGTUM, PRSARH, TBXSAG, VAHBUB, VENXIV, YUWPEL, ZIVDAJ.
17. Ramachandran, G. N. and Sasisekharan, V.,*Adv. Protein Chem*., **1968**, *23*, 283.
18. CSD refcodes: BIPVUR10, BIPWAY10, BUBLUF, BZGPRO10, CALSAR, CBBLPB10, CHPSAR, COSARC10, CPSAYL10, CTSARC, CYDSAR, CYGSGS, CYTSAR, GLSARM, HAXMOI, HXSARM, KEPNAU, LEBDEB, WEBZEI, ZAJDUJ, ZAJYUE.
19. CSD refcode: ZAJYUE; Dvorkin, A. A., Simonov, Y. A., Malinovskii, T. I., Ivanov, E. I., Kalayanov, G. D., Yaroshchenko, I. M., *Zh. Obshch. Khim.*, **1993**, *63*, 2329.
20. Coll, M., Knof, S. H., Ohga, Y., Messerschmidt, A., Huber, R., Moellering, H., Ruessmann, L. and Schumacher, G., *J. Mol. Biol.*, **1990**, *214*, 597.
21. Mikol, V., Kallen, J., Pfluegl, G. and Walkinshaw, M. D., *J. Mol. Biol.*, **1993**, *234*, 1119.
22. Mikol, V., Kallen, J. and Walkinshaw, M. D., *Protein Eng.* **1994**, *7*, 597.
23. Papageorgiou, C., Florineth, A. and Mikol, V., *J. Med.Chem.*, **1994**, *37*, 3674.
24. Mikol, V., Kallen , J. and Walkinshaw, M. D., *Proc. Natl. Acad. Sci. USA*, **1994**, *91*, 5183.
25. Altschuh, D., Vix, O., Rees, B. and Thierry, J.-C., *Science*, **1992**, *256*, 92.
26. Mikol, V., Papageorgiou, C. and Borer, X., *J. Med. Chem.*, **1995**, *38*, 3361.
27. Ke, H., Mayrose, D., Belshaw, P.J., Alberg, D. G., Schreiber, S. L., Chang, Z. Y., Etzkorn, Ho, S. and Walsh, C. T., *Structure*, **1994**, *2*, 33.
28. Ke, H., Zhao, Y., Luo, P., Weissman, I., and Friedman, J., *Proc. Natl. Acad. Sci. USA*, **1993**, *90*, 11830.
29. Maple, J. R., Dinur, U., Hagler, A. T., *Proc. Natl. Acad. Sci. USA*, 1988, *85*, 5350; Maple, J. R., Hwang, M.-J.,Stockfish, T. P., Dinur., U., Waldman, M., Ewig, C. S., Hagler, A. T., *J. Comput. Chem.*, **1994**, *15*, 162.
30. Moehle, K. and Hofmann, H.-J., *Biopolymers*, **1996**, *38*, 781.
31. Moehle, K. and Hofmann, H.-J., *J. Mol. Model*., **1996**, *2*, 307.
32. Values are given for the MP2/6-31G* calculated local minima with a *trans* amide bond at the N-terminus.
33. Allen , F. H., Harris, S. E. and Taylor, R., *J. Comput.-Aided Mol. Design*, **1996**, 10, 247.

Chapter 7

Conformational Analysis of Substance P C-Terminal Tri- and Hexapeptoids. Towards a Model for Mimicry

ABSTRACT

The conformational behaviour of substance P hexapeptoids can be understood in terms of low-energy conformations of the model peptoid monomers, Ac-NAla-NMe_2, Ac-NAla-NHMe and Ac-Gly-NMe_2, and is characterized by a near-perpendicular orientation of adjacent amide bonds in the peptoid chain. The relative directions of side chains (when viewed from the position of the backbone) in important energetically accessible peptoid conformations and a number of secondary structure elements in peptides were found to correspond. Similar observations were made for *minimum* energy conformations of tripeptide and tri-retropeptoids derived from the C-terminal sequence of substance P. This holds a promise for describing mimicry of other molecular properties by peptoids and peptides at other energetically accessible conformations as well.

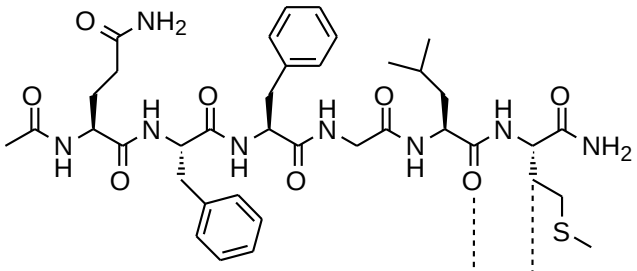
INTRODUCTION

Recently, a number of biologically active peptoid peptidomimetics has been reported. (i) Substance P (SP) peptoid analogues have been shown to be active as mimics of the undecapeptide SP at the murine NK_1 receptor; the entire SP_{1-11} peptoid and retropeptoids, as well as the hexa(retro)peptoid of the C-terminal minimal sequence of SP, i.e. SP_{6-11}, displayed full agonist activity, but with moderate potencies.[1] These peptoids have been designed based on the topological similarity of the side-chain positions in peptoids and peptides. In addition, in *retro*peptoids also the carbonyl positions with respect to the side chains are in agreement between peptoids and the parent peptide (see Fig. 1). (ii) Screening of a library of ca. 5000 di- and tripeptoids, stemming from a combinatorial synthetic approach has yielded high-affinity ligands (K_i's in the nanomolar range) for the α_1-adrenergic and μ-opiate receptors.[2] The side chains had been biased in order to resemble known ligands for a number of G-protein coupled receptors. (iii) Earlier Simon et al., reported on peptoid inhibitors for α-amylase and a hepatitis A viral protease as well as a peptoid ligand for the transactivator-responsive element RNA of HIV-1.[3] Binding data were in support of the hypothesis that *retro*peptoids are better mimics of the original peptide or protein inhibitors than the corresponding peptoids. These peptoids and retropeptoids also have been shown to posses stability towards enzymatic hydrolysis.[3]

The peptoid agonists for the murine NK_1 receptor are the first non-peptide agonists reported for any of the mammalian neurokinin receptors.[1] In addition, they are the first peptoid agonists reported constituting eleven or six peptoid (*N*-substituted glycine) residues. The SP_{6-11} hexapeptoid and retro-hexapeptoid had agonist potencies of the same order of magnitude,[1] which at first sight suggests that the peptoid backbone might be hardly of importance in interactions with the receptor. While pharmacological studies with these compounds are in progress, a number of studies has been performed to gain insight into the structural basis of the action of peptoids. The (tertiary) amide bonds in peptoids have not only been shown to allow for *cis/trans* isomerization in solution at room temperature,[4,5] which is not generally observed in peptides, but also allow for considerable pyramidalization at the amide nitrogen in both small-molecule crystal structures and in protein binding-sites (Chapter 5). Therefore, peptoids may in principle have more opportunities than peptides to comply with the requirements of the macromolecular binding-site. In addition, because of the lack of chiral centres in non-proline residues a larger part of the Ramachandran plot is in

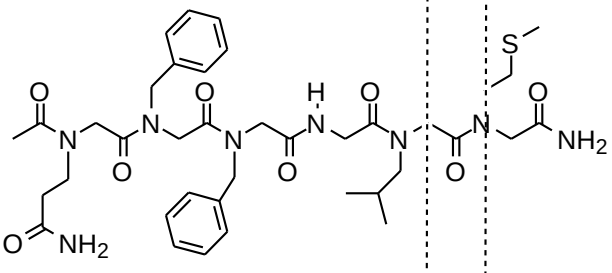
Ac-Gln^{6}-Phe^{7}-Phe^{8}-Gly^{9}-Leu^{10}-Met^{11}-NH$_2$

(**1**)



Ac-NGln^{6}-NPhe^{7}-NPhe^{8}-Gly^{9}-NLeu^{10}-NMet^{11}-NH$_2$

(**2**)



Ac-NMet^{11}-NLeu^{10}-Gly^{9}-NPhe^{8}-NPhe^{7}-NGln^{6}-NH$_2$

(**3**)

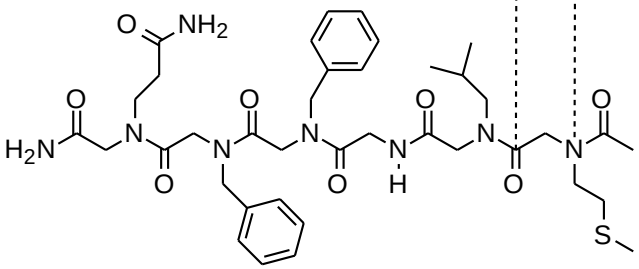


FIGURE 1. Sequences and structural formula of the C-terminal hexapeptide of substance P (**1**, SP_{6-11}), its peptoid (**2**) and retropeptoid (**3**) analogues. Residues are indexed according to their position in the amino-acid sequence of substance P.

principle available for peptoids, in comparison to peptides. Energetically accessible conformations of characteristic peptoid fragments in small-molecule crystal structures have mainly been found in the upper-left and lower-right quadrants of the Ramachandran plot

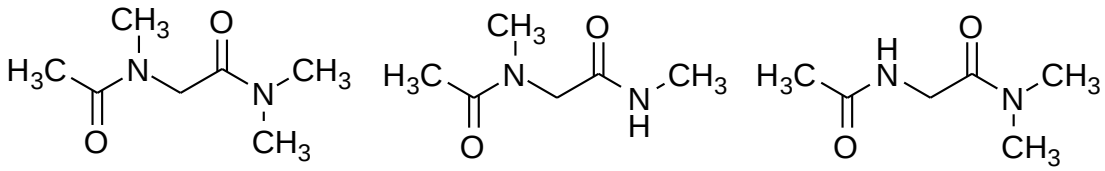
(Chapter 6). All peptoid fragment conformations could be characterized by a nearly perpendicular orientation of two adjacent amide bonds. Since experimentally determined structures of oligopeptoids are unknown to the present date, we employed theoretical conformational analysis, to study the conformational behaviour of peptoid monomers as part of peptoid hexamers and to evaluate the implications of important peptoid conformations for side-chain mimicry. The biologically active C-terminal hexapeptoid analogues of SP were studied by 1.02 ns high-temperature (1000 K) molecular dynamics simulations followed by minimization of frames from the molecular dynamics trajectory. Subsequently, the resulting conformations were compared with the conformational features of model peptoid monomers. High-temperature simulations were considered necessary to take *cis/trans* isomerization into account, in agreement with experimental findings.[4,5] Mimicry between peptide side-chains in secondary structure elements and experimentally and theoretically encountered peptoid conformations, as well as between tripeptides derived from the C-terminal sequence of SP by important and local minimum energy peptoid conformations, respectively, was also evaluated.

METHODS

Molecular models for Ac-SP_{6-11} (**1**), the hexapeptoid Ac-NGln-NPhe-NPhe-Gly-NLeu-NMet-NH_2 (**2**) and retropeptoid Ac-NMet-NLeu-Gly-NPhe-NPhe-NGln-NH_2 (**3**) and the tripeptides and tri-retropeptoids as given in Table 4 have been constructed from the standard amino-acid residue and fragment libraries in INSIGHTII[6] as well as from an in-house *N*-substituted glycine residue library within the same program. Compounds **1**-**3** were built in a fully extended conformation. Prior to further calculations all models were fully optimized; energies were minimized to a final energy gradient of 0.001 kcal·mol^{-1}·Å^{-1} using conjugate gradients. All energy calculations reported here were performed using the CFF91 force field[7] in DISCOVER2.97[8] including the CFF91 atomic charge model and a dielectric constant of 1. These optimized geometries served as starting conformations for a 1.02 ns (1020 ps) molecular dynamics simulation at 1000 K after 25 ps of equilibration. The time step in the molecular dynamics simulations was 1 fs. Atomic velocities after every time step were calculated using the leapfrog algorithm. Molecular dynamics frames were recorded every 1 ps and fully optimized according to the criteria given above.

For the model monomers Ac-NAla-NMe_2 (**4**), Ac-NAla-NHMe (**5**), Ac-Gly-NMe_2 (**6**) (Fig. 2), and the tripeptides and tri-retropeptoids energy contour plots for the ϕ and ψ

torsion angles of the central residue were constructed using an additional torsion force (a harmonic potential with a force constant of 1000 kcal·mol^{-1}·deg^{-1}) and geometry optimization of the structures; energies were minimized to a final energy gradient of 0.001 kcal·mol^{-1}·Å^{-1}. In case of the model monomers the stepsize was 10˚ and minimization was performed using conjugate gradients. A 15˚ stepsize and the VA09A(BFGS) minimizer were employed in case of the trimers. Contour plots for the tri-retropeptoids were constructed for every combination of *cis/trans* isomers along the two amide bonds, which were kept in their initial conformation by application of a extra torsion force (force constant 1000 kcal·mol^{-1}·deg^{-1}). For the tripeptides only *trans* amide bonds were considered, also with application of a torsion force. All local minimum conformations located in the trimer contour plots were additionally subjected to full geometry optimization without additional constraints.



4 **5** **6**

FIGURE 2. Structural formula of the model peptoid monomers Ac-NAla-NMe$_2$ (**4**), Ac-NAla-NHMe (**5**) and Ac-Gly-NMe$_2$ (**6**).

RESULTS AND DISCUSSION

Previously, we studied the geometry and the conformational behaviour of peptoid monomers in protein binding-sites and in the solid state (Chapters 5 and 6). In this chapter the conformational behaviour of *chains* of peptoid monomers -oligopeptoids- was studied by means of molecular dynamics and molecular mechanics simulations, using the CFF91 force field.[7] CFF91 is a generalized, second generation force field, which is parametrized to reproduce the results of *ab initio* quantum chemical calculations at the HF/6-31G* level.[7] The parametrization of peptides and peptoids in the CFF91 force field differs only for internal coordinates that are peptoid specific, *viz*. bond lengths, angles and torsion angles involving the tertiary nitrogen atom and atoms in the side chain. The so-called potential-type of the secondary and tertiary amide nitrogen atoms is, however, the same in peptides and peptoids. Auto-equivalence parameters[8] are employed for bond angles around $C_{N\alpha}$ (i.e. the

side-chain atom at the α position attached to the backbone nitrogen), torsion angles along all N-$C_{N\alpha}$ bonds, torsion angles along $C_{N\alpha}$-$C_{N\beta}$ in NPhe, NTyr, NTrp and NHis, torsion angles along $C_{N\alpha}$-$O_{N\beta}$ in NSer and NThr, and the torsion angle along $C_{N\alpha}$-$S_{N\beta}$ in NCys.

The Model Peptoid Monomer Ac-NAla-NMe_2

A comparison of quantum chemically (MP2/6-31G*) calculated minimum energy conformations[9] of the model peptoid Ac-NAla-NMe_2 (**4**) with the corresponding conformations and energies when calculated after re-optimization using CFF91 are given in Table 1. Although the CFF91 is not explicitly tailored towards the description of peptoids -or any other compounds for that matter-, there is a good agreement between both the backbone conformations and the relative energies at both levels of theory. Three independent local minima (six after including the mirror-image conformations) exist for the achiral Ac-NAla-NMe_2 model monomer, in using both CFF91 and quantum chemical calculations. The minima were found to be somewhat dependent on the presence of a *cis* or *trans* tertiary amide bond at the N-terminus, but agree very well between the two methods (Table 1). The relative energies for the *trans* amide bonds correspond well in both magnitude and order. The energy differences in case of *cis* amide bonds, as observed at the MP2/6-31G* level are, however, not reflected in the CFF91 results. Both methods, in all but one case, give higher relative energies for corresponding conformations with a *cis* than with a *trans* amide bond. The out-of-plane geometry of the tertiary amide bond is also in agreement for the two methods (see Table 3, Chapter 5). RMS differences for the non-hydrogen atoms of the conformations optimized at the MP2/6-31G* level and with CFF91 were 0.137 Å for conformation $C_{7\beta}$, 0.157 Å for α_D and 0.117 Å for α (See Table 3 for the main chain φ,ψ torsion angle combinations in these conformations). The names used to identify the conformations are the ones adopted from Moehle and Hofmann.[9]

TABLE 1. Relative energies and torsion angle values ϕ,ψ for local minimum conformations of Ac-NAla-NMe_2 (**4**) in empirical and quantum chemical geometry optimizations for both *trans* and *cis* conformations of the N-terminal amide bond

	CFF91				MP2/6-31G*[a]		
	ϕ	ψ	ΔE ($kJ \cdot mol^{-1}$)	Geometry[a]	ϕ	ψ	ΔE ($kJ \cdot mol^{-1}$)
trans	-117.5	87.1	0.0	$C_{7\beta}$	-128.2	77.0	0.0
	78.5	-166.3	4.7	α_D	74.2	-175.6	2.0
	-56.6	-52.6	15.8	α	-54.7	-47.2	24.2
cis	-134.3	76.8	8.1	$C_{7\beta}$	-153.8	62.5	14.7
	82.4	-178.2	8.0	α_D	72.4	172.1	7.9
	-62.8	-55.6	8.6	α	-62.4	-52.2	17.2

Every conformation has an equivalent mirror-image conformation, in which all torsion angle values have changed signs. a) Geometry codes by Moehle and Hofmann.[9]

The ψ-*trans* conformation (Chapter 6, and Table 2 in this Chapter) is close to the α_D minimum conformation for Ac-NAla-NMe_2. The ψ-*cis* torsion angle values as a starting conformation for Ac-NAla-NMe_2 resulted in an α conformation upon optimization using CFF91. This is in agreement with the solid state results (Chapter 6) were it has been argued that the ψ-*cis* conformation is only energetically accessible for a peptoid residue with a secondary amide or carboxylic acid group at its C-terminal side. Yet, optimization of Ac-NAla-NHMe from a ψ-*cis* starting conformation resulted in a C_7-like,[10] hydrogen bond stabilized conformation with (ϕ,ψ) values -88, 75˚. The ψ-*cis* conformation therefore does not correspond to a local minimum conformation *in vacuo*. The ψ-(+)*gauche* conformation corresponds to the global minimum ($C_{7\beta}$) conformation of Ac-NAla-NMe_2. Optimization of Ac-NAla-NMe_2 with the ϕ-(-)*gauche* starting conformation using CFF91 resulted in an α_D conformation.

Substance P C-terminal Hexapeptoid Analogues

The encouraging correspondence between the experimental, the quantum-chemical and the force-field results on the model peptoid Ac-NAla-NMe_2 prompted us to study the conformational behaviour of peptoid monomers in hexapeptoids that are based on the C-terminal sequence of substance P (SP) (see Fig. 1). The hexapeptoid **2** and hexa-retropeptoid

3 have been found to be full agonists at the (murine) NK_1 receptor, the preferred receptor for SP,[1] and are therefore of a special interest to us. In studying the conformational behaviour of peptoids *cis/trans* isomerization of the tertiary amide bonds has to be taken into account; for various peptoids, in different solvents, the presence of distinct rotamers resulting from rotation about the tertiary amide bonds has been observed in proton NMR spectra at ambient temperature.[4,5] High-temperature molecular dynamics simulations (1000 K, 1.02 ns simulation time) were performed, followed by full minimization of molecular dynamics frames taken every ps from the molecular dynamics trajectory. This resulted in 1020 local minimum energy conformations for each hexapeptoid. The equilibration time, before collecting the actual molecular dynamics trajectory, was 25 ps. Equilibration of the total energy was, however, already reached after approximately 10 ps for both hexapeptoids. The 1020 conformers in the minimized set span an energy window of 26.9 and 33.7 kcal·mol^{-1} for the hexapeptoid and the retro-hexapeptoid, respectively. The simulation temperature of 1000 K was sufficiently high to result in roughly equally occupied *cis* and *trans* conformations of the tertiary amide bonds in the minimized sets. Secondary amide bonds were mainly encountered in a trans conformation; less than 5% of the minimized conformations contained *cis* secondary amide bonds.

The backbone conformations of each residue in the sets of minimized hexapeptoids can be represented as points in a scatter diagram in ϕ,ψ space for that residue. Scatter diagrams for the residues in the hexapeptoids reveal striking similarities. Only three different ϕ,ψ scatter diagrams were discerned: (a) one corresponding to peptoid monomers connected in the chain by means of tertiary amide bonds on either end of the monomer; no appreciable differences could be observed for the different side chains (NPhe, NLeu, NMet and NGln), (b) peptoid monomers with a secondary or primary amide group at their C-terminal side, i.e. every residue succeeded by a glycine residue in the chain or every amidated C-terminal residue, (c) glycine residues. Typical scatter diagrams are given in Fig. 3, with energy contour plots for the corresponding model monomers Ac-NAla-NMe_2 (**4**), Ac-NAla-NHMe (**5**) and Ac-Gly-NMe_2 (**6**) alongside in Fig. 4. The simulation times were sufficiently long so that the scatter plots became symmetrical: all equally accessible parts of the conformational space for each residue were visited to more or less the same extent.

The occupied regions in the upper-left and lower-right part of Fig. 3 (a) correspond to low-energy conformations (0 to 2 kcal·mol^{-1}) of the model monomeric residue Ac-NAla-NMe_2 as given in Fig. 4 (a). The lower-left and upper-right quadrants correspond roughly to

the 2-4 kcal·mol^{-1} region of Ac-NAla-NMe$_2$. The distinct global minima in the Ac-NAla-NHMe energy contour plot (Fig. 4 (b)) correspond to C_7 conformations stabilized by an intramolecular hydrogen bond. This contour plot corresponds to the *trans* isomer at the N-terminus of Ac-NAla-NHMe. Since *cis* and *trans* amide bonds are present in a nearly equal amount, scatter plot Fig. 3 (b) can be regarded as a compromise of occupying the low-energy conformations of both Fig. 4 (a) and (b). In addition to the occupied regions in the scatter diagram in Fig 3(a), the scatter diagram for the glycine residues in the peptoid chains contains many conformations with either ϕ or ψ (or both) torsion angles in a *trans* conformation (Fig. 3(c)). This is in agreement with the global energy minimum conformations in the (ϕ,ψ) *trans,trans* part of conformational space for Ac-Gly-NMe$_2$. The α-helical conformations in Fig. 3(c) are only occupied in the 8-9 kcal·mol^{-1} range of the model glycine (Fig. 4(c)). In summary, local minimum conformations of oligopeptoids can be understood in terms of the low-energy conformations of these three model monomers.

The upper-left and lower-right parts of the scatterplot in Fig. 3(a) indicate conformations that clearly resemble the crystal structure conformations of representative peptoid monomers in acyclic and cyclic molecules (see Chapter 6). Also, the fairly exclusive occurrence of ψ-*cis* conformations in the peptoid monomers with a secondary amide bond at their C-terminal side (Fig. 3(b)) is in agreement with the crystallographic results. Furthermore, the occupied regions in the lower-left and the upper-right parts, with the exception of the ψ-*trans* conformations, have not been found to be occupied in the crystal structures (see Chapter 6).

Previously, Simon et al.[3] reported energy contour diagrams for Ac-Gly-NMe$_2$ and Ac-NAla-NMe$_2$ (Ac-Sar-NMe$_2$) that are qualitatively similar to the corresponding contour diagrams in Figs 3(a) and (c). In agreement with our results they suggested that the conformational behaviour of the peptoid backbone monomers (with a tertiary amide bond at their C-terminal side) is largely unaffected by the nature of the side chain and that a larger diversity of conformational states for a peptoid than for a peptide exist. Their conclusion that the α conformation, (ϕ,ψ) -55, -47°, is associated with *cis* amide bonds is, however, not in agreement with our results. The work of Simon et al., 1992, was reported while lacking any experimental structural knowledge on peptoids. That indeed the conformational behaviour of oligopeptoids could be understood in terms of the conformational behaviour of model monomers was first substantiated by our results.

A Conformational Characteristic for Peptoids

In Chapter 6, bioactive conformations of sarcosine (*N*-methylglycine) and the energetically accessible peptoid fragment conformations in the solid state as well as *in vacuo* have been characterized by the preferential perpendicular orientation of two adjacent amide bonds. This characteristic is quantified by the interplane angle θ between the least squares planes of two amide bonds. The distributions for the interplane angle θ in a peptoid residue in the minimized hexapeptoid sets were again highly similar for every non-glycine (Fig. 5 (a)) or glycine residue (Fig. 5(b). No distinction could be made between the θ distribution of a peptoid residue flanked by two tertiary amide bonds and a residue succeeded by a glycine residue in the chain. As was the case for the peptoid monomer Ac-NAla-NMe_2 *in vacuo* and the peptoid fragments in the solid state (Chapter 6), peptoid monomers in local energy minima of oligopeptoids seem to be characterized by a strong tendency towards a near-perpendicular orientation of adjacent amide bonds. A glycine residue in a peptoid chain, on the other hand, reveals partly the same behaviour, but allows for more or less coplanar orientations of two amide bonds as well. This is clearly represented by the secondary maximum at 15 to 20˚ in Fig. 5 (b). Fig. 6 shows, as a contour plot, the area of the ϕ,ψ conformational space of Ac-NAla-NMe_2 in which the interplane angle θ ranges from 70 to 90˚. Since this plot is purely geometrical (θ as a function of ϕ,ψ) it is also valid for glycine and non-proline amino-acid residues. Comparison to the corresponding energy contour plot (Fig. 4(a)) clearly reveals that for peptoid monomers θ≈90˚ corresponds to low-energy conformations in their Ramachandran plot.

The interplane angle θ does not only play an important role in local minimum conformations for peptoids, it also seems to be a characteristic of the dynamical behaviour of peptoids: in a 1 ns molecular dynamics simulation of Ac-NAla-NMe_2 at 300 K using CFF91, the interplane angle θ distribution ranges from 40 to 90˚ displaying a parabolic shape, with the majority of conformations between 70 and 90˚ and a maximum at 85-90˚ (data not shown). In a recent molecular dynamics study of Ac-NAla-NMe_2 in water at 300 K using the CHARMm22 force field, mainly conformations belonging to the α_D/ψ-*trans* region of the Ramachandran plot were present.[11] The interplane angle θ corresponding to the α_D conformation is 87˚ (see Chapter 6), which suggests that also in solution θ≈90˚ is a characteristic of energetically accessible peptoid conformations.

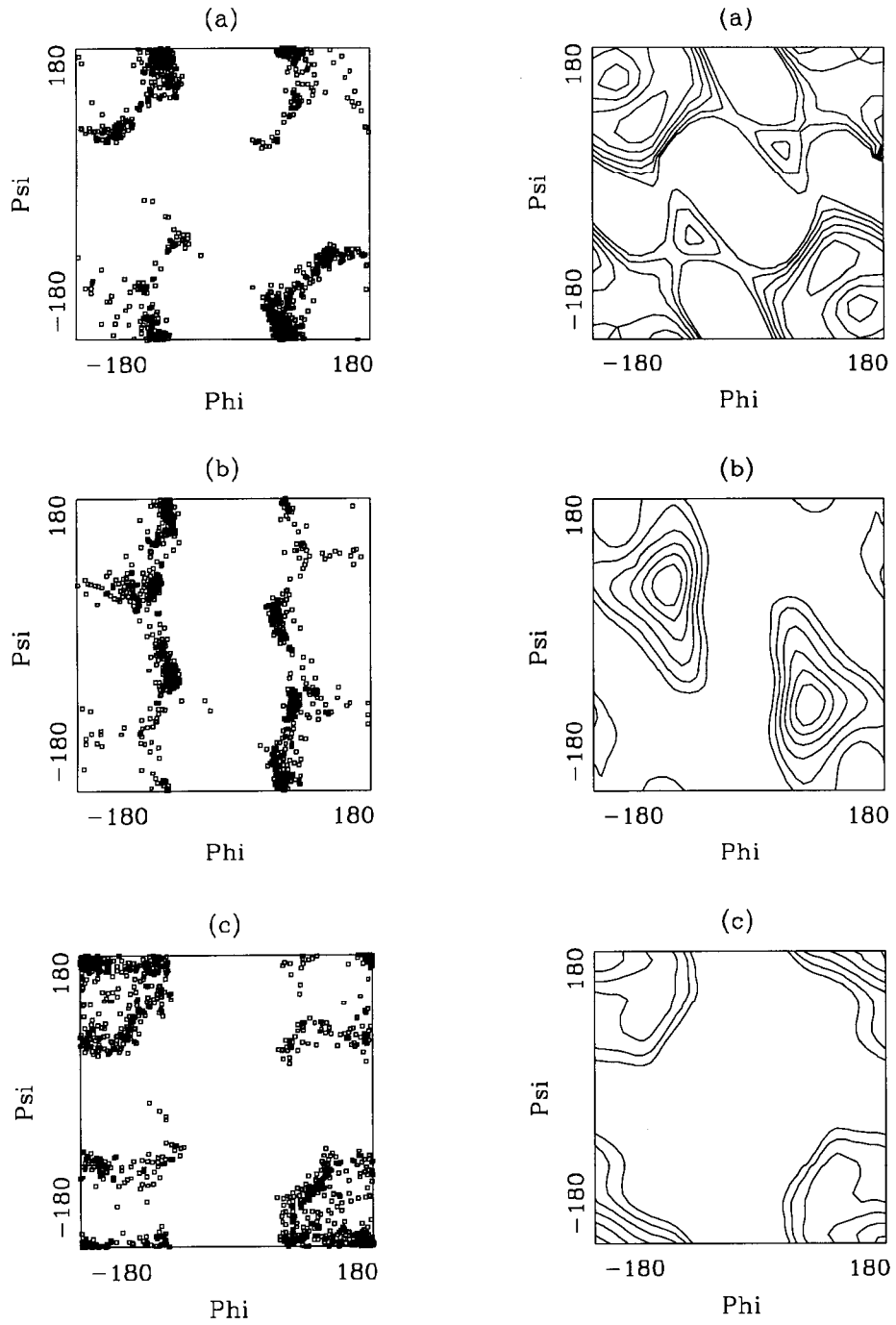


FIGURE 3 (left-hand). Typical ϕ,ψ scattergrams for peptoid residues in local minimum conformations of the peptoid and retropeptoids of $SP_{6\text{-}11}$. (a) $NPhe^{7}$ in peptoid **2**, (b) $NLeu^{10}$ in retropeptoid **3**, (c) Gly^{9} in retropeptoid **3**.

FIGURE 4 (right-hand). Energy contour plots as a function of torsion angles ϕ,ψ in model peptoid monomers (a) Ac-NAla-NMe_2, (b) Ac-NAla-NHMe and (c) Ac-Gly-NMe_2. Plots are contoured to 6.0 kcal·mol^{-1} above the plot minimum, with steps of 1.0 kcal·mol^{-1}. Global minima are indicated with a dot. In plot (c) the global minima are located at the corners of the plot.

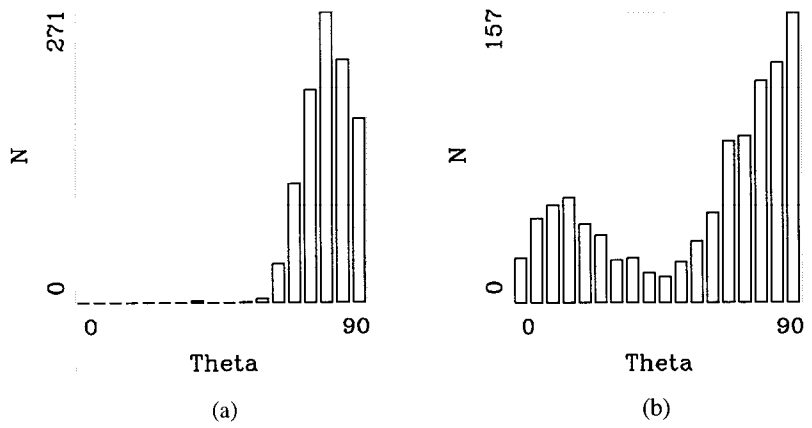


FIGURE 5. Typical interplane angle θ distribution in the local minimum conformations of the hexapeptoids. (a) NLeu^{10} in retropeptoid **3**, (c) Gly^{9} in retropeptoid **3**. Classwidths are 5˚.

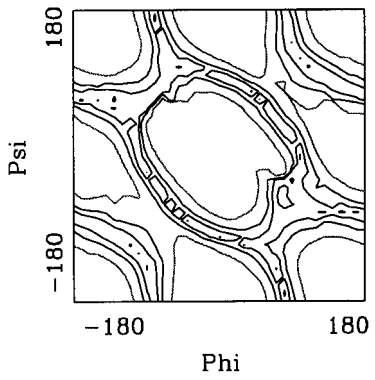


FIGURE 6. Interplane angle θ as a function of ϕ,ψ for Ac-NAla-NMe$_2$. Contoured at 70˚ (dotted line), 80, 85 and 90˚.

Mimicry of Side-chain Positions

As stated before, obtaining mimicry of the side-chain positions of a peptide in its bioactive conformation by the corresponding peptoid is the goal in designing peptoids and retropeptoids. Thus far this approach has been successful in a number of cases,[2,3] including the substance P peptoids and retropeptoids as well as the SP$_{6\text{-}11}$ hexa(retro)peptoids derived from the minimal C-terminal sequence of SP.[1] At this stage, the topological similarity of the side-chain positions in peptides and peptoids is, however, the only structure-based rationale for this approach. In peptides the relative side-chain positions in succeeding residues are determined by the ψ_i, ω_i, ϕ_{i+1} torsion angles. In peptoids they are mainly determined by the

ϕ_i, ψ_i and ω_i torsion angles (thereby neglecting the possible out-of-plane geometry of the tertiary amide bonds). The side-chain positions are for a large part determined by the positions of the C_α and C_β atoms in a peptide conformation. This finding prompted Horwell et al.[12] to exploit the direction of the C_α-C_β bonds in two adjacent amino-acid residues in a β-turn in the putative bioactive conformation of a cyclic peptide NK_2 antagonist, to design a non-peptide template which may function as a β-turn mimetic. Although the approach has not been successful in developing a non-peptide NK_2 antagonist,[12] in our opinion, it can be employed to get a first indication of the structural basis of side-chain mimicry by peptoids.

In our definition, the direction of a side chain in peptides and peptoids is given by the direction of the C_α-C_β and N-$C_{N\alpha}$ bonds, respectively. The relative direction of two side chains (not necessarily adjacent ones) is then fully determined by the (improper) dihedral τ, the angle α between the vectors indicating the directions of the bonds, and the distance *d* between the origins of the two vectors as shown in Fig. 7. Table 2 lists the relative side chain directions, as given by τ, α and *d*, for characteristic peptoid backbone conformations as encountered theoretically and experimentally (Chapter 6), as well as for important secondary structure elements in peptides.[13,14] In reading Table 2, it has to be taken into account that the mirror-image conformations of peptoids are also possible and are energetically equivalent, this in contrast to the peptide conformations. In addition, in accordance with experimental findings,[4,5] *cis/trans* isomers of peptoids were explicitly included.

Correspondence between the side chain directions in these peptoid and peptide conformations is indicated in Table 3. Initial correspondence with the side-chain directions for two adjacent residues in periodic secondary structure elements in peptides was in some cases, however, lost, because the corresponding peptoid conformation can not exist as a periodic structure (Table 3). Of the peptoid/peptide conformations with positive correspondence only the ψ-*trans* peptoid conformations were sterically possible as periodic structures. Note that in a few cases corresponding side-chain directions are only observed due to the existence of *cis* amide bonds in peptoids. In conclusion, Table 3 indicates that peptoid side-chain mimicry is possible for a number of secondary structure elements in peptides, but presents no proof of that: the entire molecule must be taken into account to assess the possibilities of mimicry. Molecular modelling is in that respect a necessary tool to propose energetically accessible conformations of both peptoids and peptides, since these are not restricted to the ones given in Table 2, and to evaluate the ability of functional groups to be involved in intermolecular interactions.

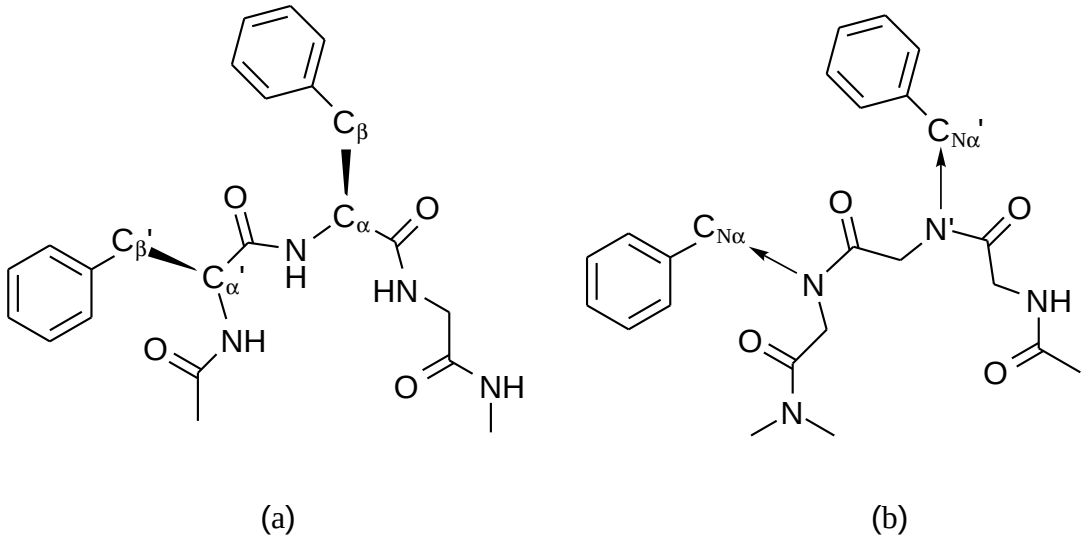


FIGURE 7. The relative direction of a side chain with respect to the other (and *vice versa*) in (a) a given peptide conformation is completely described by the improper dihedral $\tau(C_\beta'\text{-}C_\alpha'\text{-}C_\alpha\text{-}C_\beta)$, the angle α between the vectors $\mathbf{v}(C_\alpha'\rightarrow C_\beta')$ and $\mathbf{w}(C_\alpha\rightarrow C_\beta)$ and the distance $d(C_\alpha'\text{-}C_\alpha)$. (b) In peptoids the analogous definitions are: the improper dihedral $\tau(C_{N\alpha}'\text{-}N'\text{-}N\text{-}C_{N\alpha})$, the angle α between the vectors $\mathbf{v}(N'\rightarrow C_{N\alpha}')$ and $\mathbf{w}(N\rightarrow C_{N\alpha})$ and the distance $d(N'\text{-}N)$.

To further assess the usefulness of the relative side-chain directions in studying mimicry of peptides by peptoids, we compared local minimum conformations of tripeptides and corresponding tri-retropeptoids from the C-terminal sequence of SP. The trimers were derived by passing a three-residue "window" along Gly^9 of SP, which is an important residue in determining the subtype selectivity of SP for the NK_1 and NK_2 receptor.[15] The sequences are given in Table 4. From the non-exhaustive set of local minimum conformations of the tripeptides and tri-retropeptoids, as derived from the ϕ,ψ-energy contour plot of the central residue in every trimer (see Methods), the pairs best mimicking each others side-chain directions were selected (see Fig. 8); conformations were superimposed based on the corresponding atom pairs of the atoms included in the definitions for τ, α and d (Fig. 7). The backbone conformations are given in Table 4. The three peptide conformations are partially stabilized by intramolecular hydrogen bonds forming C_7 conformations,[10] whereas in the tri-retropeptoids intramolecular hydrogen bonds are absent. Meanwhile, the interplane angles θ range from 76 to 90˚. Residue Gly^9 in retropeptoid **8** (Table 4) is an exception with nearly coplanar amide bonds (θ=8˚). In Figs. 8 (a) and (b), where the glycine residues are termini of the trimers, the backbone and the first side-chain atoms of the non-glycine residues

TABLE 3. Correspondence between side-chain directions in preferential peptoid conformations (horizontal) and secondary structure elements in peptides (vertical)

Peptoids	$C_{7\beta}$/ ψ-(+)*gauche*		ϕ-(-)*gauche*		α_D/ψ-trans		α	
Peptides	*cis*	*trans*	*cis*	*trans*	*cis*	*trans*	*cis*	*trans*
α-helix			(x)		x			
3_{10}-helix			(x)		x			
parallel β-sheet							(x?)	
antiparallel β-sheet							(x?)	
extended							(x?)	
poly(L-proline) I								
poly(L-proline) II						x		
type I/type I' β-turn								
type II/type II' β-turn								x?
γ-turn		x		x		x		

The letter x indicates correspondence in τ, α and *d* as given in Table 2. A question mark indicates doubtful correspondence. Parentheses indicate that the associated peptoid conformations are sterically incompatible with a periodic structure.

roughly occupy the same space. The end-groups of the residues do not generally match in these conformations. The glycine residues are totally misaligned, so in fact a superposition of dimers seems to have been performed. Note, however, that the glycine residues do play an important role in stabilizing the peptide conformations. In Fig. 8(b) where glycine is the central residue, both the trimer backbones as well as the side chains roughly occupy the same space. This correspondence was obtained for minimum energy conformations, but only the similarity of the occupied volumes is considered. Although these findings can not explain the experimentally observed mimicry of the SP peptoids,[1] they exemplify that mimicry obtained at this level (occupied volumes in local minimum conformations) holds a promise for explaining mimicry of many other molecular properties when considering all energetically accessible conformations. These properties can for instance be solvent accessibilities of groups, the molecular electrostatic potential, or a combination of properties combined in so-called molecular fields.[16,17] For exploiting these findings in ligand design molecular modelling techniques are indispensable. Given the conformational flexibility of both peptides and peptoids, it seems plausible that conformationally constrained peptoid analogues are necessary to direct the search for the bioactive conformation of SP and to improve selectivity, affinity and agonist activity of the peptoids in the future. Combined efforts in this respect from the fields of molecular modelling, structural chemistry and organic synthesis are also deemed necessary in the future.

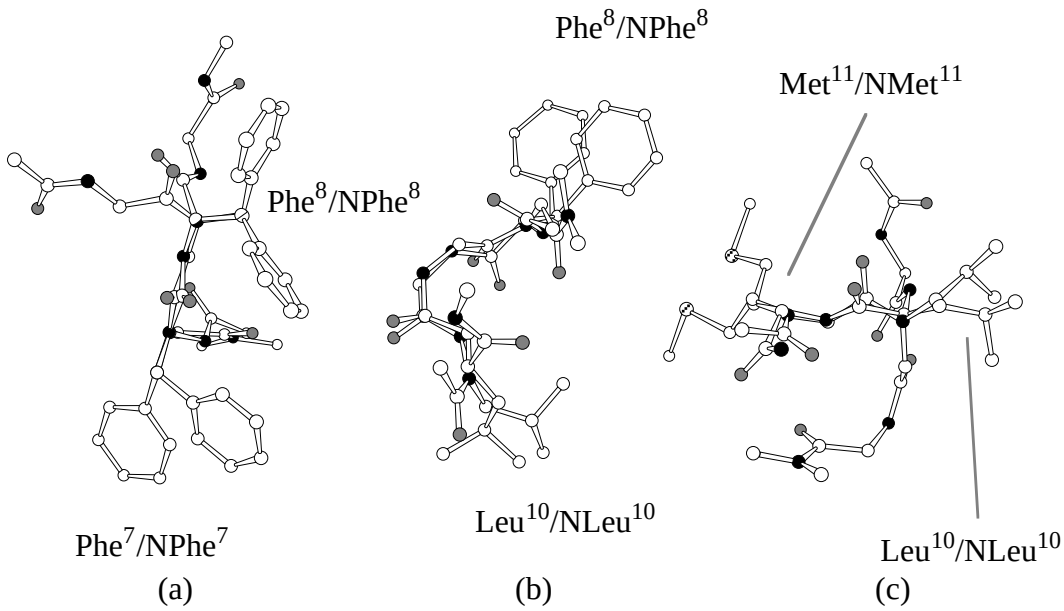


FIGURE 8. Superpositions of local minimum conformations of the tripeptides and corresponding tri-retropeptoids with the largest similarity of the side chain positions considering the direction of the bonds C_α-C_β and N-$C_{N\alpha}$. (a) **7** and **8**, (b) **9** and **10**, (c) **11** and **12**. Backbone conformations are given in Table 2.

TABLE 4. Conformations of the best matching tripeptide and tri-retropeptoid local minimum conformations

No	Sequence	ϕ_i	ψ_i	ω_i	ϕ_{i+1}	ψ_{i+1}	ω_{i+1}	ϕ_{i+2}	ψ_{i+2}	ω_{i+2}
7	Ac-Phe^{7}Phe^{8}Gly^9-NHMe	-84	72	*-178*	-165	-61	*178*	83	-70	*180*
8	Ac-Gly^{9}NPhe^{8}NPhe^7-NMe_2	177	179	*-8*	-77	-174	*2.5*	77	176	-
9	Ac-Phe^{8}Gly^{9}Leu^{10}-NHMe	-90	84	*-174*	89	-59	*178*	-87	75	*-176*
10	Ac-NLeu^{10}Gly^{9}NPhe^8-NMe_2	-83	180	*7*	77	-166	*-1*	77	-175	-
11	Ac-Gly^{9}Leu^{10}Met^{11}-NH_2	82	-68	*-177*	-84	66	*-175*	-90	74	-
12	Ac-NMet^{11}NLeu^{10}Gly^9-NMe_2	-70	131	*-9*	67	-159	*20*	84	-163	-

Torsion angles that determine the side-chain directions have been framed; the residues of the side chains involved have been underlined. Amide bond torsion angles are in italics.

ACKNOWLEDGEMENT

We thank Roeland Boer for his work on the peptoid and peptide trimers and Dr Ed Moret for critically reading the manuscript. The basis of the work presented here was established during a stay at the Theoretical Medicinal Chemistry group of the Janssen Research Foundation, Beerse, Belgium. We wish to express our gratitude to the Janssen Research Foundation for this opportunity.

REFERENCES

1. Westra-de Vlieger, J. F., Van Heuven-Nolsen, D, Wilting, J., Kruijtzer, J. A. W., Liskamp, R. M. J. and Nijkamp, F. P., *in preparation*, **1997**.
2. Zuckermann, R. N., Martin, E. J., Spellmeyer, D. C., Stauber, G. B., Shoemaker, K. R., Kerr, J. M., Figliozzi, G. M., Goff, D. A., Siani, M. A., Simon, R. J., Banville, S. C., Brown, E.G., Wang, L., Richter, L. S. and Moos, W. H., *J. Med. Chem.*, **1994**, *37*, 2678.
3. Simon, R. J., Kania, R. S., Zuckermann, R. N., Huebner, V. D., Jewell, D. A., Banville, S., Ng, S., Wang, L., Rosenberg, S., Marlowe, C. K., Spellmeyer, D. C., Tan, R., Frenkel, A. D., Santi, D. V., Cohen, F. E. and Bartlett, P. A., *Proc. Natl. Acad. Sci. USA*, **1992**, *89*, 9367.
4. Kruijtzer, J. A. W. and Liskamp , R. M. J., *Tetrahedron Lett.*, **1995**, *36*, 6969.
5. Boks, G. J., Boer, D. R., Kruijtzer, J. A. W., Liskamp, R. M. J., Tollenaere, J. P., Kroon, J. and Schouten, A., *unpublished results*, **1997**.
6. Biosym/MSI, 9685 Scranton Road, San Diego, CA 92121-2777 USA, 1995.
7. Maple, J. R., Dinur, U., Hagler, A. T., *Proc. Natl. Acad. Sci. USA*, 1988, *85*, 5350; Maple, J. R., Hwang, M.-J.,Stockfish, T. P., Dinur., U., Waldman, M., Ewig, C. S., Hagler, A. T., *J. Comput. Chem.*, **1994**, *15*, 162.
8. DISCOVER User Guide, Part 1, October 1995, San Diego, Biosym/MSI, 1995.
9. Moehle, K. and Hofmann, H.-J., *Biopolymers*, **1996**, *38*, 781.
10. Rommel-Moehle, K. and Hofmann, H.-J., *J. Mol. Struct.*, **1993**, *285*, 211.
11. Moehle, K. and Hofmann, H.-J., *J. Mol. Model*., **1996**, *2*, 307.
12. Horwell, D. C., Howson, W., Naylor, D. and Willems, H. M. G., *Bioorg. Med. Chem. Lett.*, **1995**, *5*, 1445.
13. IUPAC-IUB Commission on Biochemical Nomenclature, *J. Mol. Biol.*, **1970**, *52*, 1.
14. Liskamp, R. M. J., *Recl. Trav. Chim. Pays-Bas*, **1994**, *113*, 1.
15. Ward, P., Ewan, G. B., Jordan. C. C., Ireland, S. J., Hagan, R. M. and Brown, J. R., *J. Med. Chem.*, **1990**, *33*, 1848.
16. Vinter, J. G. and Trollope, K. I., *J. Comput.-Aided Mol. Design*, **1995**, *9*, 297.
17. Chapter 3 of this thesis.

TABLE 2. Relative side chain directions of adjacent residues in characteristic peptoid conformations and in (idealized) secondary structure elements[13,14] in peptides. Backbone torsion angles in degrees

Peptoid[a]	ϕ_i	ψ_i	ω_i	τ (°)	α (°)	d (Å)
Theoretical						
$C_7\beta$	-128	77	-176	138	128	3.1
	-128	77	-7.2	16	78	3.1
α_D	74	-176	174	128	130	3.7
	74	-176	13	-74	94	3.7
α	-55	-47	172	37	41	3.0
	-55	-47	-4	150	154	3.0
Experimental						
ψ-(+)*gauche*	-128	67	180	124	118	3.0
	-128	67	0	-1	72	3.0
ϕ-(-)*gauche*	-69	141	180	-120	114	3.6
	-69	141	0	78	120	3.6
ψ-*trans*	-85	175	180	-105	108	3.7
	-85	175	0	77	106	3.7
ψ-*cis*	-88	2	180	63	60	2.6
	-88	2	0	53	101	2.6

Peptide	ϕ	ψ	ω	τ (°)	α (°)	d (Å)
Periodic structures:						
α helix (right-handed)	-57	-47	180	75	83	3.8
3_{10} helix	-60	-30	180	87	96	3.8
parallel β-sheet	-119	113	180	177	178	3.8
antiparallel β-sheet	-139	135	180	179	178	3.8
extended chain	180	180	180	180	180	3.8
poly(L-proline) I	-83	158	0	42	78	2.8
poly(L-proline) II	-78	149	180	-103	109	3.8

reverse	ϕ_{i+1}	ψ_{i+1}	ϕ_{i+2}	ψ_{i+2}	τ (°)	α (°)	d (Å)
(β) turns:[b]							
type I	-60	-30	-90	0	53	69	3.8
type I'	+60	+30	+90	0	-62	70	3.8
type II	-60	+120	+80	0	31	45	3.8
type II'	+60	-120	-80	0	-20	35	3.8
γ-turn:[b]	ψ_i	ϕ_{i+1}	ψ_{i+1}	ϕ_{i+2}	τ (°)	α (°)	d (Å)
	120	-65	80	-120	-116	125	3.8

In the peptoid and periodic peptide conformations and the γ-turn, the relative side chain directions correspond to residues i and i+1, while in case of the β-turns the side chain in residues i+1 and i+2 are described. a) *Theoretical*: values for MP2/6-31G* optimized geometries of Ac-NAla-NMe_2 with a *trans* N-terminal amide bond taken from Moehle and Hofmann.[9] The side-chain orientation parameters for the C-terminal *trans* or *cis* amide bond have been

calculated for the corresponding methyl group of the dimethylamide end-group. *Experimental*: see Chapter 6. b) The ω_i, ω_{i+1} and in case of the β-turns also the ω_{i+2} amide bonds need to be in a *trans* conformation in order to facilitate the stabilizing hydrogen bond.

Chapter 8

General Discussion

A relatively small number of biologically active peptoids have been reported to this date (see Chapter 1). Any consideration with respect to their mechanism of action was restricted to the 2D similarity between the peptoids and the endogenous ligand, which, as has been shown for the α_1-adrenergic receptor, needs not even to be a peptide.[1] Of the peptides reported, the substance P (SP) peptoids may be considered the most interesting. They are the only agonist peptoids reported so far that were constructed by a straightforward translation of the parent peptide sequence as well as of the minimal peptide sequence $SP_{6\text{-}11}$.[2,3] In addition, only proteinogenic side chains are present in these peptoids. From the point of view of the use of peptoid peptidomimetics in a general ligand design strategy this is of considerable interest. The SP peptoids act as full agonists for the murine NK_1 receptor.[2,3] They are, however, of a considerable lower potency than SP or $SP_{6\text{-}11}$ themselves. The selectivity and affinity of these peptoids for the neurokinin receptors remains to be established, yet the concept of peptidomimetics has proven its value. To start unravelling the mechanism of action of peptoids in general, two aspects need to be addressed. The first one concerns the structure and conformational behaviour of peptoids as a compound class. The second aspect deals with the mimicry of peptides by their corresponding peptoids. Both aspects were addressed in this thesis with reference to ligands for the (human) NK_1 receptor.

Structural chemical knowledge on peptoids is relatively scarce and especially experimental 3D structural knowledge has been essentially lacking. The covalent (2D) structure of the SP peptoids has been established by means of mass spectrometry.[3] In addition *cis/trans* isomerization of the tertiary amide bonds in SP related peptoids has been revealed by ^{1}H NMR spectra in different solvents at ambient temperature.[4] Yet, we would also like to find answers to the following questions. What do we know about (1) the structure and (2) energetically accessible conformations of the backbone in peptoid oligomers? (3) What is the influence of the various side chains on this conformational behaviour? In other words, can something be said about the sequence dependence of the peptoid conformational behaviour? (4) In comparison to peptides, what will be the consequences of the loss of hydrogen bond donors and chiral centres in the peptoid backbone? (5) Are there any stabilizing interactions or preferential folds in relation to peptoids, such as e.g. present in

peptide conformations stabilized by intramolecular hydrogen bonds? And finally, (6) what are the molecular implications of the increased flexibility as reflected in the *cis/trans* isomerization of the tertiary amide bonds? Tentative rules for the description of peptoid conformations were devised to improve communication with respect to these issues (see Appendix A). The IUPAC-IUB nomenclature for amino acids and peptides thereby served as a basis.

The average amide bond geometries in Appendix B as determined from crystal structures in the Cambridge Structural Database reflect the structure of a peptoid backbone in a crystallographically determined protein-peptoid complex, with the exception of the backbone torsion angles and the bond angle around C_α. The intrinsic out-of-plane coordination geometry of the tertiary amide nitrogen as revealed by quantum-chemical calculations (Chapter 5) was put in a slightly different perspective by recent gas electron diffraction experiments on *N*,*N*-dimethylacetamide.[5] Depending on the exact method employed to interpret these recent results, the coordination geometry at the tertiary amide nitrogen can be considered to possess either a non-planar, planar or pseudo-planar equilibrium structure *in vacuo*. The observed flexibility at the amide nitrogen atom is on the other hand in good agreement with our results (Chapter 5). Pyramidalization at the tertiary amide nitrogen in peptoids in complex with a receptor is not *a priori* necessary, but may well be an additional possibility to optimize complementarity between the peptoid and the protein. This flexibility might also imply that peptoids are less sensitive to stereochemical (chiral) requirements in a binding pocket.

The limited diversity in the sarcosine (*N*-methylglycine) binding sites prevents unambiguous conclusions with respect to the conformational behaviour of this peptoid monomer in protein binding sites (Chapter 6). Yet, an important characteristic that was observed in both peptoid monomer fragments in small-molecule crystal structures (Chapter 6) as well as in theoretical conformational analyses of e.g. peptoid oligomers (Chapter 7), was also present in these protein-ligand complexes: two adjacent tertiary amide bonds prefer a mutually perpendicular orientation. Both in static structures as well as in dynamical situations the methylene group bridging two adjacent amide bonds appears to act as a ball-joint at which a near-perpendicular orientation is preferentially maintained. It is expected that this characteristic feature is also of importance in bioactive conformations of peptoid oligomers. In addition, we hypothesize that this characteristic plays a similar role in

stabilizing peptoid conformations, as is the presence of hydrogen bonds in stabilizing secondary structural elements in peptides and proteins.

In the theoretical conformational analysis (Chapter 7) of the peptoid monomer Ac-NRA-NMe_2 our results were found to be in agreement with high-quality quantum chemical calculations by Moehle and Hofmann.[6] In addition, it was revealed that local minimum energy conformations of the SP hexapeptoids can be understood in terms of low-energy conformations (0-4 kcal/mol in the force field applied) of three isolated model *monomers*. The suggestion made by Simon et al.[7] that the conformational behaviour of the peptoid backbone is largely unaffected by the nature of the side chains is substantiated, but also refined by our results. It appears that a non-proline containing peptoid sequence can conformationally be broken down in (a) residues with tertiary amide bonds at either end, (b) residues that are succeeded by a glycine residue in the chain or that terminate in a primary (or secondary) amide group, and (c) glycine residues. Conformational flexibility of these residues increases in corresponding order. Although the conformational preferences of proline and glycine residues that are part of ligands in protein binding sites are yet to be determined (see Chapter 6), some sequence dependence of the conformational behaviour of peptoid backbones is thus already revealed by our calculations. In addition, L-Pro and D-Pro residues may be employed to restrict the monomer backbone to well defined values in the (upper)left and the (lower)right part of the Ramachandran plot, respectively. With reference to the structure-activity relationship of $SP_{6\text{-}11}$ (Chapter 1), incorporation of proline residues at position 9 in the SP peptoids seems an interesting step.

An important step in studying the conformational and interaction behaviour of peptoids at protein binding sites would be the crystallographic determination of protein-peptoid complexes. In addition, theoretical docking studies of peptoids (or peptides containing peptoid residues) to macromolecular structures for which the specific peptoids or peptides are established ligands, are considered highly valuable. Furthermore, the determination of non-complexed peptoid crystal structures will not only be of relevance for experimentally determined peptoid *oligomer* structures, but in view of the results in Chapters 2 and 3 also for studying the interactions of peptoids with their environment. The crystallization of linear, unrestrained peptoids oligomers is, however, cumbersome and difficult to achieve (Chapter 4).

The second aspect in unravelling the mechanism of action of peptoids is mimicry. The non-peptide NK_1 antagonists rather than SP itself played an important role in studying molecular recognition. Results from the antagonist crystal structures together with calculated interaction geometries (Chapters 2 and 3) suggest that indeed aromatic-aromatic, N^+-aromatic and Coulombic interactions involving formally charged groups, simultaneously can occur between quinuclidine NK_1 antagonists and their environment. This is in good agreement with the aromatic and the (perhaps) formally charged amino-acid residues (Glu-193, Lys-194 and His-197 in TM-V and His-265, Phe-268 and Tyr-272 in TM-VI) that have been reported to mainly form the antagonist binding site.[8] The likely interaction site (Chapter 2) for the side chain of Gln-165 (TM-IV) that is probably involved in a hydrogen bond with the exocyclic nitrogen or oxygen atom of the quinuclidine and piperidine antagonists may in future studies be combined with these results. The results presented and methods employed in this thesis may then finally be combined in e.g. a four-helical (TM-III, IV, V, VI) binding site model of the human NK_1 receptor. Since the bioactive conformations of quinuclidine as well as piperidine NK_1 antagonists have not been established, different antagonist conformations may be subjected to XED dockings with e.g. benzene, an imidazole ring or a primary amide probe to arrive at a consistent molecular picture for structure and interactions.

In a simplified model to study peptide and peptoid mimicry involving relative side-chain directions, indications for explaining mimicry were obtained (Chapter 7). The side-chain directions associated with the energetically accessible peptoid monomer conformations correspond to the side-chain directions of adjacent amino-acid residues in a number of secondary structure elements in peptides. Similarity, or better, mimicry at this level was also obtained for local minimum energy conformations of tripeptoids and tripeptides. This similarity is not sufficient to explain mimicry of peptoids and peptides, but revealed that *cis* amide bonds may play an important role in obtaining mimicry. Given the larger diversity of conformational states for peptoid monomers when compared to peptides, it is probably necessary to conformationally bias the peptoid building blocks in a combinatorial compound library. Proline residues are an obvious choice when a functional side chain is not required; e.g. position Gly^9 in the SP hexapeptoids is then a likely candidate (*vide infra*). The modelled peptoid helices in Chapter 6 suggest that methylating the C_α position in peptoid monomers might oppose the formation of *cis* amide bonds and thereby also influence electrostatic properties (e.g. the dipole moment). In addition, results from Simon at al.[7] suggest that depending on the chirality at this C_α carbon the monomer backbone

conformation might be restricted to the upper-left ("L-Ala") or the lower-right ("D-Ala") quadrant of the Ramachandran plot. In conclusion, such conformationally biased building blocks might be useful for increasing the agonist potency of new SP peptoid analogues and for the elucidation of the mechanism of action of SP.

REFERENCES

1. Zuckermann, R. N., Martin, E. J., Spellmeyer, D. C., Stauber, G. B., Shoemaker, K. R., Kerr, J. M., Figliozzi, G. M., Goff, D. A., Siani, M. A., Simon, R. J., Banville, S. C., Brown, E.G., Wang, L., Richter, L. S. and Moos, W. H., *J. Med. Chem.*, **1994**, *37*, 2678.
2. Westra-de Vlieger, J. F., Van Heuven-Nolsen, D, Wilting, J., Kruijtzer, J. A. W., Liskamp, R. M. J. and Nijkamp, F. P., *in preparation*, **1997**.
3. Kruijtzer, J. A. W., *Synthesis of Peptoid Peptidomimetics*, thesis, Utrecht University, 1996.
4. Kruijtzer, J. A. W. and Liskamp , R. M. J., *Tetrahedron Lett.*, **1995**, *36*, 6969.
5. Mack, H.-G. and Oberhammer, H., *J. Am. Chem. Soc.*, **1997**, *119*, 3567.
6. Moehle, K. and Hofmann, H.-J., *Biopolymers*, **1996**, *38*, 781-790.
7. Simon, R. J., Kania, R. S., Zuckermann, R. N., Huebner, V. D., Jewell, D. A., Banville, S., Ng, S., Wang, L., Rosenberg, S., Marlowe, C. K., Spellmeyer, D. C., Tan, R., Frenkel, A. D., Santi, D. V., Cohen, F. E. and Bartlett, P. A., *Proc. Natl. Acad. Sci. USA*, **1992**, *89*, 9367.
8. Gether, U., Lowe, J. A., III and Schwartz, T. W., *Biochem. Soc. Trans.*, **1995**, *23*, 96.

Tentative Nomenclature for the Description of the Conformation of Peptoids

1. Since the topology of peptoids is similar to that of peptides the IUB-IUPAC nomenclature for amino acids and peptides[1,2] can serve as a basis for the nomenclature of peptoids. Meanwhile, the practical usage of the abbreviations and symbols suggested, has prevailed any attempt to be exhaustive.

1a. In cases where for technical reasons, e.g. computer outputs, the use of Greek characters is not supported, the following Roman equivalents are recommended: α, A; β, B; γ, G; δ, D; ε, E; ζ, Z; η, H; τ, T; υ, U; ϕ, F; χ, X; ψ, Q; ω, W.[2]

2. **Peptoid monomeric building blocks** are formally *N*-alkylated glycine amino acids. The substituent on the amino nitrogen atom is called the side chain. Side chains may be identical to proteinogenic amino-acid side chains, but need not necessarily be so. Here, only substituents derived from proteinogenic amino-acid side chains are considered.

2a. Systematic names of building blocks. Based on rule 3AA-7:[1] *N*-alkylglycine, etc. Proline and glycine are peptoid building blocks by themselves and hence remain called proline and glycine.

3. A **peptoid** is any compound produced by the formation of an amide bond between the carboxylic group of one building block and the amino group of another. The amide bonds will be tertiary or secondary (in the case of glycine) amide bonds.

3a. The **direction of the main chain** in peptoid oligomers is given from the N-terminus to the C-terminus unless explicitly stated otherwise.

4. *Short-hand notation for residues*

4a. Naming of residues in a **'three/four-letter notation'**: Xaa for an amino-acid residue, NXaa for the corresponding peptoid residue (ref. 3, p.11). Exceptions are proline and glycine: Pro and Gly, respectively.

4b. **Chain starters and terminators**. A free amino group at the N-terminal is (optionally) denoted with an H-, an acetyl group is denoted with Ac. A free acid on the C terminus is (optionally) denoted -OH. Amidation of the C-terminal end can e.g. be a primary amide ($-NH_2$), methylamide (-NHMe) or dimethylamide ($-NMe_2$).

e.g.:

Substance P: H-Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH_2

peptoid: H-NArg-Pro-NLys-Pro-NGln-NGln-NPhe-NPhe-Gly-NLeu-NMet-NH_2

4c. **One-letter notation**: A lowercase n precedes the one-letter notation for the corresponding amino-acid (ref. 3, p.77). Proline and glycine are given by their amino-acid one-letter notation.

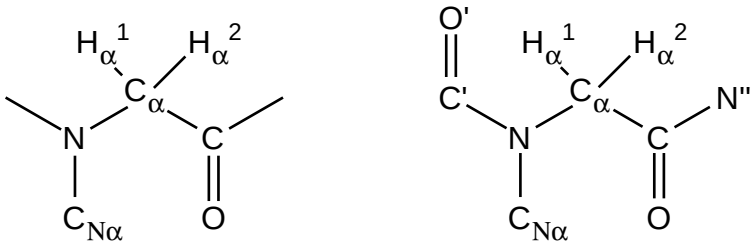
e.g.:

Substance P peptoid: N-nRPnKPnQnQnFnFGnLnM-NH_2

4d. Various *molecular modelling programmes* apply a strict three-letter notation only e.g. in the construction of fragment libraries. A peptoid residue is then given as NRX in which X is the one-letter notation of the corresponding amino-acid.

5. *Atom names in peptoid residues in chains.*

Nomenclature for **backbone atoms** in amino acids and peptides is largely maintained:[2], (Rule 1)



5a. When atoms belong to a different residue, this can be indicated by a prime or double prime. When the position *i* of a residue in an oligomer is of importance this can be indicated as $C_\alpha(i)$.

5b. H_α^1 is assigned to the H giving the smallest clockwise passage to H_α^2 when viewed in the direction of the main chain, i.e. in the C_α-C direction.[2]

5c. $C_{N\alpha}$ is the **first side-chain atom** attached to the amide nitrogen N; i.e. at the α-position with respect to the amide nitrogen. When $C_{N\alpha}$ is the only non-hydrogen side chain atom present, such as in *N*-methylglycine (sarcosine, NAla) or is used as the 'first side-chain atom' in a general sense, the name C_N is more appropriate.

5d. **Subsequent side-chain atoms** (non-hydrogens) follow the order of the characters in the Greek alphabet that replace the subscript α in $C_{N\alpha}$, corresponding to their relative positions with respect to the backbone nitrogen atom N. In the case of branching the Greek character is

directly followed by a figure (number) ; these figures follow the (priority) rules for peptide side chains.[2], (Rule 4.3)

5e. Hydrogen atoms in the side chain are designated with the same index (Greek letter and number) of the atom to which they are attached. When two or three hydrogen atoms are attached to the same atom, an additional figure (1, 2 or 3) is added in agreement with the corresponding peptide rules.[2], (Rule 4.4)

6. *Backbone (main chain) torsion angles*

Definitions for **backbone torsion angles** are:

$\phi(i)=(C(i-1)-N(i)-C_\alpha(i)-C(i))$ or $\phi(i)=(C'-N-C_\alpha(i)-C)$ or $\phi=(C'-N-C_\alpha-C)$

$\psi(i)=(N(i)-C_\alpha(i)-C(i)-N(i+1))$ or $\psi(i)=(N-C_\alpha(i)-C-N'')$ or $\psi=(N-C_\alpha-C-N'')$

$\omega(i)=(C_\alpha(i)-C(i)-N(i+1)-C_\alpha(i+1))$ or $\omega(i)=(C_\alpha(i)-C-N''-C_\alpha'')$ or $\omega=(C_\alpha-C-N''-C_\alpha'')$

6a. The level of detail is selected in correspondence with the requirements of the description of the system.

6b. In a *cis* amide bond $-90<\omega(i)<90^\circ$, in a *trans* amide bond $90<\omega(i)<-90^\circ$.

7. *Amide bond out-of-plane parameters*

7a. The four torsion angles describing the **conformation of the amide bond** are:

$\omega1=(C_\alpha-C-N''-C_\alpha'')$ or $\omega_1(i)=(C_\alpha(i)-C(i)-N(i+1)-C_\alpha(i+1))$

$\omega2=(O-C-N''-C_{N\alpha}'')$ or $\omega_2(i)=(O(i)-C(i)-N(i+1)-C_{N\alpha}(i+1))$

$\omega3=(O-C-N''-C_\alpha'')$ or $\omega_3(i)=(O(i)-C(i)-N(i+1)-C_\alpha(i+1))$

$\omega4=(C_\alpha-C-N''-C_{N\alpha}'')$ or $\omega_4(i)=(C_\alpha(i)-C(i)-N(i+1)-C_{N\alpha}(i+1))$

7b. In the case of a secondary amide bond $C_{N\alpha}$ is replaced by H or H_N.

7c. Torsion angle $\omega1$ corresponds to the backbone torsion angle ω of the amide bond at the C-terminal side of the corresponding residue. The amide bond torsion angles are related by the condition $(\omega1+\omega2)-(\omega3+\omega4)=0 \bmod 2\pi$.

7d. The **out-of-plane geometry** of amide bonds can be completely described in terms of three parameters:[4] χ_C, χ_N and τ, which quantify the non-planarity of the coordination geometry at the amide carbon atom, at the amide nitrogen atom and the 'twisting' of these two along the amide bond, respectively. Definitions are as follows:

$\chi_C = (\omega1-\omega3+\pi) \bmod 2\pi$	
$\chi_N = (\omega2-\omega3+\pi) \bmod 2\pi$	
$\tau = (\omega1+\omega2)/2$	in case of a *cis* amide bond.
$\tau = (\omega1'+\omega2')/2 \bmod 2\pi$	in case of a *trans* amide bond, ω1 and ω2 should be taken in a 0-360˚ interval, which is denoted by the primes.

7e. A completely planar coordination geometry is characterized by $\chi_C = 0˚$, $\chi_N = 0˚$ and $\tau = 180˚$ in a *trans* and 0˚ in a *cis* amide bond. A full pyramidalization at the amide nitrogen ('sp3-geometry') corresponds to $\chi_N = 60˚$.

8. *Side chains.*

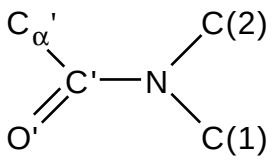
8a. **Bonds** are designated by means of the two interconnecting atoms, e.g. N-$C_{N\alpha}$. Bond identifiers to be used as superscripts for indicating torsion angles (Rule 8b) are one or more numbers (figures) separated by commas, 1 for bond N-$C_{N\alpha}$, and e.g. in the case of NVal: 2,1 for $C_{N\alpha}$-$C_{N\beta1}$ and 2,2 for $C_{N\alpha}$-$C_{N\beta2}$.

8b. **Side chain torsion angles** are denoted by χ (or χ_N if confusion arises) and one, two (or three) superscripts, separated by commas, indicating the bond. The atoms included in the torsion angle definition of a particular bond are given by the (priority) rules for atoms.[2] For instance, in peptoids torsion angle χ^1 is given by C'-N-$C_{N\alpha}$-$C_{N\beta}$, when no ambiguity for the last atom exists.

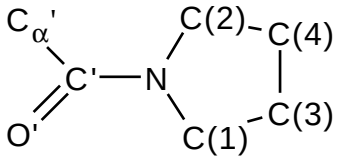
REFERENCES

1. IUPAC Joint Commission on Biochemical Nomenclature, Nomenclature and Symbolism for Amino Acids and Peptides, Recommendations 1982, *J. Biol. Chem.*, **1985**, *260*, 14.
2. IUPAC-IUB Commission on Biochemical Nomenclature, Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains, Tentative Rules 1969, *J. Mol. Biol.*, **1970**, *52*, 1.
3. Kruijtzer, J. A. W., *Synthesis of Peptoid Peptidomimetics*, thesis, Utrecht University, 1996.
4. Dunitz, J. D. and Winkler, F. K., *Acta Cryst.*, **1975**, *B31*, 251; Winkler, F. K. and Dunitz, J. D., *J. Mol. Biol.*, **1971**, *59*, 169.

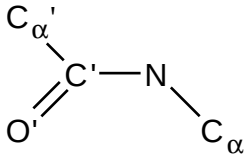
Average Solid-State Amide-Bond Geometries



1



2



3

TABLE. Average tertiary and secondary amide bond geometries in crystal structures

	1 n=104		**2** n=108		**3** n=46	
Bonds	Length (Å)	σ	Length (Å)	σ	Length (Å)	σ
C_α'-C'	1.526	0.022	1.531	0.021	1.515	0.021
C'-O'	1.231	0.013	1.233	0.012	1.230	0.013
C'-N	1.341	0.017	1.341	0.015	1.331	0.016
N-C(1)	1.469	0.018	1.468	0.013	-	-
N-C(2)	1.464	0.013	1.474	0.014	-	-
N-C_α	-	-	-	-	1.453	0.014
C(1)-C(3)	-	-	1.531	0.018	-	-
C(3)-C(4)	-	-	1.489	0.049	-	-
C(4)-C(2)	-	-	1.502	0.036	-	-
Angles	**Angle (°)**	**σ**	**Angle (°)**	**σ**	**Angle (°)**	**σ**
C_α'-C'-N	118.6	1.7	119.2	2.2	116.0	1.2
C_α'-C'-O'	119.6	1.4	120.2	1.5	120.9	1.6
O'-C'-N	121.8	1.5	120.5	1.7	123.1	1.3
C'-N-C(1)	118.4	1.3	119.2	1.6	-	-
C'-N-C(2)	124.8	1.8	128.6	2.2	-	-
C(1)-N-C(2)	116.6	1.3	111.8	1.2	-	-
C'-N-C_α	-	-	-	-	122.5	2.0
N-C(1)-C(3)	-	-	103.4	1.0	-	-
N-C(2)-C(4)	-	-	103.1	1.3	-	-
C(1)-C(3)-C(4)	-	-	104.7	1.9	-	-
C(3)-C(4)-C(2)	-	-	106.0	3.6	-	-

C(1) is the carbon atom more or less eclipsed with respect to O' (O'-C'-N-C(1) within -90 and 90°). In a *trans* amide bond along the main chain C(1) corresponds to C_α.

Summary

Developing peptoid peptidomimetics is part of a strategy to come to a more rational development of ligands for therapeutically interesting macromolecular targets for which peptide-protein or protein-protein interactions play a role. While maintaining the functionalities and relative side-chain positions of the parent peptide, peptidomimetics are likely to have improved pharmacokinetics in comparison with peptides. In addition, the modular build-up of peptoids (oligomers of *N*-substituted glycine residues) facilitates the synthesis of compounds targeted at a plethora of biological systems. To validly implement this strategy for many potentially interesting targets, insight in the structure and conformational behaviour of peptoids, which forms the topic of this thesis, is of importance.

Here, the system of interest is the tachykinergic system with the undecapeptide substance P as its most prominent member. Substance P is a neuropeptide that is released from sensory nerve endings throughout the body and is involved in e.g. neurogenic inflammation and in the transmission of pain. As such, substance P plays a role in allergic airway diseases, such as allergic asthma, and in chronic pain. Antagonists of substance P can therefore perhaps be used as analgesics or anti-inflammatory compounds. Recently, peptoid analogues of substance P have been found to elicit full agonist activity at the (murine) NK_1 receptor, which is the preferred receptor for substance P. The explanation for their activity is, however, still based on the topological similarity between the structural formula of the peptoids and the parent peptide. To gain insight in their molecular mechanism of action, the structure and conformational behaviour of peptoid peptidomimetics is studied in this thesis. Furthermore, intermolecular interactions in the crystal structures of non-peptide antagonists for the human NK_1 receptor are described and serve as a basis for evaluating a method which can be employed in the study of mimicry.

In the absence of an experimentally determined NK_1 receptor structure, or antagonist-receptor complex, the interaction geometries between antagonists and amino-acid residues that are important in binding are yet unknown. In Chapter 2, crystal structures of nine non-peptide NK_1 antagonists are analyzed for the intermolecular interactions of their pharmacophoric groups with neighbouring molecules in the crystals. Based on these analyses we identified several N^+-aromatic and aromatic-aromatic interaction geometries for the positively charged quinuclidine rings and the pharmacophoric aromatic rings, respectively, which can explain the importance of aromatic amino-acid residues in the human NK_1

antagonist binding-site. In addition, an interaction site for Gln-165 in the human NK_1 receptor is explicitly proposed. This amino-acid residue has been reported to be involved in a hydrogen bond with the benzylamino nitrogen or benzylether oxygen of piperidine and quinuclidine based antagonists. A superposition of the crystal structure conformations of two prototypic NK_1 antagonists, CP-96,345 and CP-99,994, based on pharmacophoric elements, revealed similarities in intermolecular interaction geometries in the crystal environment, that might be used as a starting point to explain the similar binding-sites at the human NK_1 receptor that have been indicated for these two compounds.

The number and types of interactions encountered in crystal structures of low-molecular-weight compounds are severely restricted by the limited functionality present in the crystal, especially when compared to the functionality present in protein binding-sites. Also, although the crystal structure conformation of a compound is an energetically accessible conformation, it is in many cases not likely to be the bioactive one. To avoid these drawbacks the applicability of an empirical method developed to calculate favourable interaction sites and interaction geometries, irrespective of the molecular conformation or the nature of the interacting groups, was evaluated for a selection of non-peptide NK_1 antagonists. This is reported in Chapter 3. Experimentally determined interaction sites for hydrogen bond acceptor groups and N^+-aromatic and aromatic-aromatic interaction geometries (as reported in Chapter 2) could be reproduced remarkably well. This method would facilitate the development of a pharmacophore based on interaction sites, and the construction of a model for the NK_1 antagonist binding-site. Application of this method would also be highly valuable in studying mimicry of peptoids and peptides.

Experiences with crystallization attempts of substance P and its (retro)peptoid peptidomimetic are described in Chapter 4. Crystal structures of substance P and its retropeptoid would contribute to the fundamental understanding of their energetically accessible conformations and their intermolecular interactions in a highly structured molecular environment. There are, however, no precedents with respect to successful crystallizations of linear and otherwise unrestrained peptide oligomers of a comparable size. The crystallization strategy employed was based on the use of macromolecular crystallization screens in combination with a hanging drop-vapour diffusion set-up. For both substance P and its retropeptoid no initial crystallization conditions were identified. Crystallization of

substance P was prevented by the formation of phase separations and amorphous precipitates, which may be related to the amphiphilic nature of substance P.

The consequences of the presence of tertiary amide bonds for the backbone geometry of peptoids, with emphasis on the geometry of the amide bonds itself, are described in Chapter 5. In *in vacuo* quantum chemical calculations tertiary amide bonds were found to be intrinsically non-planar. In the solid state, on the other hand, a preferred planar geometry was observed that still allows for considerable pyramidalization at the amide nitrogen in individual cases. In ligands in protein binding-sites a similar degree of pyramidalization at the tertiary amide nitrogen was observed. In *trans* amide bonds the monomer side chain is mainly responsible for a non-planar amide bond, whereas in *cis* amide bonds the peptoid backbone takes over this role. Pyramidalization at the tertiary amide nitrogen seems to be an additional feature for peptoids in accommodating their conformation to the requirements of a protein binding-site.

Chapter 6 turns to the relatively flexible parts of the peptoid backbone. The energetically accessible conformations for a fragment characteristic of the conformationally flexible methylene bridge connecting the relatively rigid amide bonds in a peptoid backbone were evaluated in small-molecule crystal structures and in ligands in crystallographically determined protein complexes. The methylene group was found to act as a ball-joint at which the amide bonds prefer a mutually perpendicular orientation. The most prominently present peptoid monomer conformation is shown to be compatible with the formation of a periodic structure.

Conformational analysis of tri- and hexapeptoid analogues derived from the C-terminal sequence of substance P, as described in Chapter 7, reveals that a mutually perpendicular orientation of adjacent amide bonds is also a characteristic of peptoid monomers conformations when they are part of a peptoid chain. The conformational behaviour of the peptoid backbone can be understood in terms of low-energy conformations of the model peptoid monomers Ac-NAla-NMe_2, Ac-NAla-NHMe and Ac-Gly-NMe_2. These model compounds represent in corresponding order (i) any peptoid monomer that is part of the peptoid chain by means of tertiary amide bonds on either end, (ii) a peptoid monomer that precedes a glycine residue in the chain and (iii) a glycine residue. In addition, the relative directions of side chains (when viewed from the position of the backbone) in prominent

energetically accessible peptoid conformations and a number of secondary structure elements in peptides were found to correspond. Similar observations were made for *minimum* energy conformations of tripeptide and tri-retropeptoids derived from substance P. Mimicry between peptides and peptoids is found at this level, but is not sufficient to explain the biological activity of the substance P peptoids or any other peptoid. Though, it is a starting point for describing mimicry of other relevant molecular properties in energetically accessible peptoid and peptide conformations as well.

The discussion in Chapter 8 focuses on the insight gained in the structure and conformational behaviour of peptoid peptidomimetics as a result of the work described in this thesis. In summary, some idea about the characteristics and the sequence dependence of the backbone conformational behaviour of peptoids has emerged. These features also seem to describe the behaviour of peptoids in protein binding-sites. Molecular modelling studies aimed at directing synthetic efforts and to interpret pharmacological structure-activity studies with respect to the substance P peptoids and other peptoid peptidomimetics can be aided and validated by the methods and results presented in this thesis. It may form the start of a more rational development of peptidomimetic ligands for therapeutically interesting targets. Final validation can of course only be obtained by experiment.

Samenvatting

Het ontwikkelen van peptoïde-peptidomimetica is deel van een strategie die streeft naar een rationelere ontwikkeling van liganden voor biologische macromoleculen, waarbij peptide-eiwit of eiwit-eiwit interacties een rol spelen, en die ook mogelijk therapeutisch van belang zijn. Peptidomimetica hebben naar verwachting betere farmacokinetische eigenschappen dan het oorspronkelijke peptide, terwijl functionele groepen en de onderlinge zijketenposities uit het peptide behouden blijven. Daar komt nog bij dat de modulaire opbouw van peptoïden (oligomeren van *N*-gesubstitueerde glycine residuen) de synthese van verbindingen voor talloze biologische systemen mogelijk maakt. Om deze strategie echter zinvol toe te passen op de vele mogelijk interessante targets is inzicht nodig in de structuur en het conformationeel gedrag van peptoïden. Dit proefschrift is gewijd aan het verkrijgen van dat inzicht.

Het systeem dat hier centraal staat is het tachykinerge systeem, met het undecapeptide substance P als de meest belangrijke vertegenwoordiger. Substance P is een neuropeptide dat wordt vrijgemaakt uit de uiteinden van sensorische zenuwen die door het hele lichaam aanwezig zijn. Het is betrokken in ontstekingsprocessen die aangestuurd worden vanuit het zenuwstelsel en speelt een rol in de overdracht van pijnprikkels. Dit speelt bijvoorbeeld een rol bij allergische asthma, of andere allergische aandoeningen in de luchtwegen, en bij chronische pijnen. Antagonisten van substance P kunnen dus wellicht werkzaam zijn als analgetica of als ontstekingsremmende middelen. Recent bleek dat peptoïde analoga van substance P volledige agonistische activiteit op de NK_1 receptor van muizen vertoonden. De NK_1 (neurokinine-1) receptor is het receptor-eiwit waar substance P bij voorkeur aan bindt. Als verklaring voor deze activiteit kan echter alleen nog maar de topologische overeenstemming in de structuurformules van de peptoïden en het oorspronkelijke peptide worden gegeven. Om inzicht te krijgen in hun werkingsmechanisme worden in dit proefschrift de structuur en het conformationeel gedrag van peptoïden onderzocht. Tevens worden kristalstructuren van niet-peptide antagonisten voor de humane NK_1 receptor bestudeerd en hun intermoleculaire interacties beschreven. Deze dienen vervolgens als basis voor het evalueren van een methode die gebruikt kan worden om mimicry verder te bestuderen.

Omdat een experimentele structuur van de NK_1 receptor, of van de receptor gecomplexeerd met een antagonist nog ontbreekt, zijn ook de interactie-geometrieën tussen

antagonisten en de bindende aminozuren nog onbekend. In Hoofstuk 2 worden de intermoleculaire interacties van farmacofore groepen met de hen omringende moleculen bestudeerd zoals ze aanwezig zijn in kristalstructuren van negen niet-peptide NK_1 antagonisten. Op basis van deze resultaten konden we verschillende interactie-geometrieën identificeren, zoals N^+-aromaat interacties voor de positief geladen quinuclidine ringen en aromaat-aromaat interacties voor farmacofore aromaatringen, die het belang van aromatische aminozuurresiduen in de antagonist bindingsplaats in de humane NK_1 receptor kunnen verklaren. Tevens wordt expliciet een interactieplaats voor Gln-165 uit de humane NK_1 receptor voorgesteld. In de literatuur wordt gerapporteerd dat dit aminozuurresidue een waterstofbrug vormt met het stikstofatoom van de benzylamino groep, of het zuurstofatoom van de benzylethergroep in piperidine en quinuclidine NK_1 antagonisten. Superpositie van de kristalstructuurconformaties van twee representatieve NK_1 antagonisten op basis van hun farmacofore groepen bracht overeenkomsten in de intermoleculaire interactiegeometrieën in de kristalomgeving aan het licht. Deze overeenstemming kan wellicht als uitgangspunt dienen voor het verklaren van de bevinding dat deze twee verbindingen op eenzelfde plaats aan de humane NK_1 receptor binden.

Het aantal en het type van de interacties die waargenomen worden in kristalstructuren van verbindingen met een laag moleculair gewicht worden erg ingeperkt door de beperkte functionaliteit die aanwezig is in het kristal, zeker als dit vergeleken wordt met de functionaliteit in bindingsplaatsen van eiwitten. Daarbij geldt ook, dat ook al is de kristalstructuurconformatie van een verbinding een energetisch toegankelijke conformatie, het in veel gevallen niet waarschijnlijk is dat dit ook de biologisch actieve conformatie is. Om deze nadelen te omzeilen werd een op empirie gebaseerde methode geëvalueerd, die ontwikkeld is om gunstige interactieplaatsen en interactiegeometrieën te kunnen berekenen, onafhankelijk van de moleculaire conformatie en de aard van de interagerende groepen. De methode werd toegepast op een selectie van niet-peptide NK_1 antagonisten. Dit wordt beschreven in Hoofdstuk 3. Het bleek dat zowel experimenteel bepaalde interactieplaatsen voor waterstofbrug accepterende groepen als N^+-aromaat en aromaat-aromaat interactiegeometrieën (zoals beschreven zijn in Hoofdstuk 2) vermeldenswaardig goed konden worden gereproduceerd. Deze methode zou het mogelijk kunnen maken een farmacofoor te ontwikkelen die gebaseerd is op interactieplaatsen en tevens behulpzaam kunnen zijn in het maken van een model voor de NK_1 antagonist bindingsplaats. Gebruik van deze methode in het bestuderen van de mimicry van peptoïden en peptiden is ten zeerste aan te raden.

Ervaringen die zijn opgedaan tijdens kristallisatiepogingen met substance P en het bijbehorende retropeptoïde peptidomimeticum worden beschreven in Hoofdstuk 4. De kristalstructuren van substance P en zijn retropeptoïde zouden bijdragen aan het fundamentele begrip van zowel hun energetisch toegankelijke conformaties als hun intermoleculaire interacties in een gestructureerde moleculaire omgeving. Er bestaan echter geen precedenten voor succesvolle kristallisaties van lineaire peptiden met een vergelijkbare grootte, die niet op enige wijze conformationeel beperkt zijn. De kristallisatiestrategie werd gebaseerd op het gebruik van een selectie macromoleculaire kristallisatiecondities in combinatie met een hangende druppel-gas/vloeistof kristallisatie methode. Voor substance P noch voor het retropeptoïde werden initiële kristallisatiecondities gevonden. Kristallisatie van substance P werd verhinderd door het optreden van faseschеidingen en de vorming van amorfe neerslagen, dat waarschijnlijk veroorzaakt wordt door het amfifiele karakter van substance P.

De gevolgen van de aanwezigheid van tertiaire amidebindingen voor de geometrie van de peptoïde backbone worden, met nadruk op de geometrie van de amide bindingen zelf, beschreven in Hoofdstuk 5. Tertiaire amidebindingen bleken in quantum chemische berekeningen *in vacuo* intrinsiek (d.w.z. zonder omgevingsinvloeden) niet-planair. In de vaste fase daarentegen, werd een bij voorkeur planaire geometrie waargenomen, waarbij echter in individuele gevallen een nadrukkelijk pyramidale omringing van het amide stikstofatoom mogelijk blijft. Bij liganden in bindingsplaatsen van eiwitten werd een overeenkomstige mate van pyramidalisatie van het stikstofatoom in de tertiaire amidebinding waargenomen. In *trans* amidebindingen is de zijketen van het monomeer met name verantwoordelijk voor een niet-planaire amidebinding, terwijl bij *cis* amidebindingen de peptoïde backbone deze rol vervult. Een pyramidale omringing van het stikstofatoom in tertiaire amidebindingen lijkt voor peptoïden een aanvullende mogelijkheid te zijn om hun conformatie aan te passen aan de vereisten van een bindingplaats in een eiwit.

Hoofdstuk 6 richt zich op de relatief flexibele gedeeltes van de peptoïde backbone. In klein-molecuul kristalstructuren en in liganden uit kristallografische bepaalde eiwit-ligand complexen werd nagegaan wat de energetisch toegankelijke conformaties zijn van het conformationeel flexibele deel van de peptoïdebackbone, te weten de methyleen brug die twee relatief starre amidebindingen met elkaar verbindt. De methyleen groep bleek als kogelgewricht te werken waarbij de amide bindingen een onderling loodrechte stand

verkiezen. De peptoïde-monomeerconformatie die het meest nadrukkelijk optreedt is verenigbaar met de vorming van een periodieke structuur uitgaande van deze conformatie.

In Hoofdstuk 7 wordt een conformatieanalyse beschreven van tri- en hexapeptoïde analoga die zijn afgeleid van de C-terminale sequentie van substance P. Deze analyse toont aan dat een onderling loodrechte stand van naburige amidebindingen ook karakteristiek is voor de conformaties van peptoïdemonomeren die deel uitmaken van een peptoïdeketen. Het conformationeel gedrag van de peptoïdebackbone kan worden geïnterpreteerd in termen van de lage-energieconformaties van de model-peptoïdemonomeren Ac-NAla-NMe_2, Ac-NAla-NHMe en Ac-Gly-NMe_2. Deze verbindingen staan model voor achtereenvolgens (i) ieder peptoïdemonomeer dat via twee tertiaire amidebindingen in de peptoïdeketen is opgenomen, (ii) een peptoïdemonomeer dat in de keten gevolgd wordt door een glycine residue en (iii) een glycine residue zelf. Verder bleek dat de relatieve richtingen van zijketens (gezien vanuit de backbone) in belangrijke, energetisch toegankelijke peptoïdeconformaties met een aantal secundaire structuurelementen in peptiden overeen te komen. Soortgelijke bevindingen werden gedaan voor minimumenergieconformaties van, uit substance P afgeleide, tripeptiden en triretropeptoïden. Mimicry van peptiden door peptoïden op dit niveau is niet voldoende om de biologische activiteit van de substance P peptoïden, of enig ander peptoïde, te verklaren. Het is echter een startpunt om ook mimicry van andere van belang zijnde moleculaire eigenschappen in energetisch toegankelijke peptoïde- en peptideconformaties te beschrijven.

De discussie in Hoofdstuk 8 richt zich op het inzicht, dat door het in dit proefschrift beschreven werk is verkregen in de structuur en het conformationeel gedrag van peptoïde-peptidomimetica. Samengevat, er is een indruk verkregen van de karakteristieken en de sequentie-afhankelijkheid van het conformationeel gedrag van de peptoïdebackbone. Deze kenmerken lijken eveneens het gedrag van peptoïden in bindingsplaatsen van eitwitten te beschrijven. De methoden en resultaten die gepresenteerd zijn in dit proefschrift kunnen in belangrijke mate bijdragen aan het uitvoeren van valide molecular modelling studies die gericht zijn op de sturing van synthese-inspanningen en de interpretatie van farmacologische structuur-activiteitsrelaties met betrekking tot de substance P peptoïden en andere peptoïde-peptidomimetica. Het kan de opmaat vormen voor een meer gestuurde ontwikkeling van peptidomimetische liganden voor therapeutisch interessante targets. De uiteindelijke validatie kan uiteraard alleen uit experimenten worden verkregen.

Dankwoord

Jan Tollenaere, ik heb je als promotor en vriend leren kennen in de discussies die we gevoerd hebben over allerhande wetenschappelijke en veelal ook niet-wetenschappelijke onderwerpen. Magda, jij hebt hier ook een belangrijke en altijd gastvrije rol in gespeeld. Jan Kroon, eveneens als promotor, ik ken niemand met een zo'n grote kristallografische kennis en kunde. Jij bent me bijna altijd te slim af. En zo hoort het ook, zou je zeggen. Jan Tollenaere en Jan Kroon, het was een voorrecht om met zulke ervaren wetenschappers te werken. Vooral de laatste drukke maanden heb ik veel van jullie uithoudingsvermogen gevraagd, maar ik denk dat we het er gezamelijk niet onaardig vanaf hebben gebracht. Ik heb er tot het laatste hoofdstuk plezier in gehouden. Ik ben jullie zeer erkentelijk.

Ed Moret, je bent de meest constante factor geweest tijdens mijn promotieonderzoek en je hebt veel meer aan de richting en de vorm van mijn project bijgedragen dan je je waarschijnlijk zelf realiseert. Bedankt hiervoor. Wij zijn allebei verschrikkelijk eigenwijs, maar weten daar gelukkig constructief mee om te gaan. Ik heb er grote bewondering voor hoe je de dagelijkse leiding van de CMC groep vorm hebt weten te geven. Uitbreiding van de groep is van groot belang. Of ook Alex in die lange-termijnplanning is opgenomen weet ik niet, maar Inge-en-een-half en jij hebben de smaak wel te pakken.
Hans Hilbers, jij zorgde voor leven in de brouwerij en was altijd direct bereid om problemen en probleempjes op te lossen. Niemand is onmisbaar, maar zonder jou zou de zaak toch een stuk saaier zijn geweest. Andy Vinter, part-time but most honorary member of our group. Keep up the good work!
Met de modelling 'aio's' was het aanvankelijk droevig gesteld, maar de Wetenschap trekt. Alfred Kostense, jij liet de apotheek deels zijn beloop en ging iets echt interessants doen; het beste bewijs van je motivatie. Malcolm Gillies en Rachel, overgevlogen from 'down under' en vervolgens bemerken dat het Nederlandse vreemdelingenbeleid nog wel wat tekortkomingen kent. Igor Schillevoort, die vanaf zijn tweedejaars praktikum al zijn kunnen en ambities mbt onderzoek heeft getoond en nu met zijn eigen promotieonderzoek aan de slag kan. Jullie hebben de noodzakelijke kritische massa danig vergroot. Ik ben blij dat het zo gegaan is.

Janneke Westra-de Vlieger, helaas hebben onze projecten te weinig synergie kunnen vertonen. Dat doet overigens niets af aan onze prettige persoonlijke samenwerking. Ik heb er grote bewondering voor hoe je je werk bij farmacologie en medicinal chemistry/farmaco-chemie hebt verdeeld en hoe je met de problemen die opdoken tijdens je werk bent omgegaan. De komst van Lisette is waarschijnlijk het beste geweest dat je tot nu toe overkomen is. Wij hebben alleen nog maar twee katten, maar gaan door met oefenen. Ik wens je met Roel en Lisette nog veel geluk toe en succes met de afronding van je boekje.
Marcel Fischer, ja, ik ben te weinig op de zesde geweest, maar de keren dat ik er was waren zeer de moeite waard. Als je je nieuwe onderzoek net zo vurig ter hand neemt als onze discussies, wordt het zeker een succes. Jeannette Paulussen, samen met Janneke en Natatia binnen het 'mannenbolwerkje' farmacochemie. Je was een fijne en gezellige collega en ik ben blij dat je ondanks je interesse voor een andere richting toch je onderzoek met een promotie hebt afgesloten (kijk eens naar *Science*, 269, 903.) Marc Maliepaard, voor eeuwig zal jouw naam verbonden zijn aan het Legendarische Eerste Kerstdiner in 1993. Natatia Lemmert, Nico de Mol en Bert Janssen, ieder voor zijn eigen inbreng tijdens mijn promotietijd bedankt.

Maar laten we eens bij het begin beginnen. Er was eens...Jaap Wilting, jij was mijn eerste kennismaking met Farmacie. Pas later kon ik bevroeden wat dat voor consequenties

had. De tweede persoon die ik binnen de faculteit leerde kennen was Linda Hutzezon. Sommige mensen vergeet je nooit meer...

In februari 1993, nog voor mijn project begonnen was, werd een organisch chemicus uit Leiden geïntroduceerd, die wellicht, in de toekomst, wat interessante analoga van substance P zou kunnen leveren... Binnen een paar jaar werd echter de bijbehorende groep uit Leiden naar Utrecht getransplanteerd. Rob Liskamp, jij was eerder hoogleraar dan ik gepromoveerd! Jij hebt het stimulerende enthousiasme voor onderzoek weer teruggebracht in onze gelederen. Zonder jouw komst zou dit boekje er heel anders hebben uitgezien. John Kruijtzer, Dennis Löwik, Karen Sliedregt-Bol, Suzanne Mulders, Astrid Boeijen, John Reichwein, Rob Ruytenbeek, Dries de Bont, Arwin Brouwer en natuurlijk Lovina Hofmeyer bedankt voor het mede maken van wat Medicinal Chemistry nu is!

Een behoorlijk deel van mijn tijd heb ik doorgebracht bij de eiwitkristallografie groep. Dat betekende dat ik weer aan het pipetteren werd gezet bij het opzetten van kristallisatie experimenten. Jolanda Hurenkamp, jij hebt mij in het begin veel geholpen en de handigheid bijgebracht om 'druppels op te hangen'. Dat experimentele werk en de empirische aanpak zijn erg nutig geweest om de rest van mijn theoretische werk beter te organiseren.
Piet Gros, bedankt voor de gelegenheid die je me geboden hebt om met niet-macromoleculaire verbindingen aan de slag te gaan en toch de eiwitkristallografie heb kunnen verkennen. De wekelijkse werkbesprekingen en de theorieavonden met de groep van Titia van het NKI waren erg stimulerend. Arie Schouten, als praktische vraagbaak en allround kristallograaf ben je een niet weg te denken factor in mijn werk geweest. Juist die combinatie van kristallografische kennis mbt laag-moleculaire en macromoleculaire verbindingen is denk ik een belangrijke meerwaarde van de groep. In dat verband wil ik ook Ton Spek en Toine Schreurs bedanken voor hun medewerking.
Op de eiwit-aio's ben ik altijd een beetje jaloers geweest, want hoe interessant het theoretische werk aan biomoleculen ook is, ook de experimentele kant blijft trekken. Jeroen Brandsen, Carien Dekker, Raimond Ravelli en Barend Bouma, maar ook de niet-aio's Boris Strokopytov, Erik Huizinga, Andrew McCarthy en Jean van den Elsen, jullie waren een leuke groep om mee te werken.

Na ongeveer zes weken (ongeveer) bij Farmacie werd ik meteen opgezadeld met het illustere studententrio Harald van den Born, Reinout Schellekens en Jasper van Gorp. Na een stoomcursus modelling zijn jullie in de VS in het diepe gesprongen en hebben het er fantastisch van afgebracht. Alle bewondering hiervoor. Al snel volgden twee jonge dames, Muck van Groningen en Irma Rigter. Jullie deden een moeilijk project in Cambridge bij Andy Vinter. Ik heb jullie daar even opgezocht en, niet verwondelijk, daar zijn de niet-wetenschappelijke zaken me beter van bijgebleven dan de rest. Enige tijd later begon Jacco Pippel met een bijvak. Binnen de kortste keren wist je meer 'trif-trafjes' van de unix systemen dan onze systeembeheerder. Als hekkesluiter voor de verandering eens een scheikunde student, Roeland Boer. Achter je af en toe wat 'flegmatische houding' (geleend citaat) bleek de ware onderzoeksdrift boven te komen. We hebben het aan jou te danken dat we nu over de kristalstructuur van een peptoïde beschikken. Een goede keuze van je om verder te gaan in het onderzoek. Veel succes.

In de loop van de vier jaren heb ik de warme belangstelling van mijn ouders, schoonouders, grootouders, broer en schoonzus erg op prijs gesteld. Allemaal bedankt hiervoor, ook al was jullie waarschijnlijk niet altijd helemaal duidelijk wat ik precies uitspookte. Ducky en Sammy en jullie bijna continue eetlust mogen in dit lijstje ook niet ontbreken. Prrrr.
Angelique, jouw voortdurende liefde en steun waren onontbeerlijk voor het voltooien van deze etappe. Laten er nog vele gelukkige volgen. Ik hou van je.

CURRICULUM VITAE

Gertjan Boks werd geboren op 21 augustus 1969 te Elst (Gld.). In 1987 behaalde hij het Atheneum diploma aan het Van Lingen College te Arnhem. In 1988 behaalde hij het propedeutisch examen en in februari 1993 het doctoraal examen scheikunde aan de Universiteit Utrecht. Bij het Research Instituut voor Toxicologie (RITOX) (Prof. dr W. Seinen, Utrecht) heeft hij een bijvak Biochemische Toxicologie volbracht. Zijn extra keuzevak was Theoretische Chemie bij de vakgroep Theoretische Chemie (Prof. dr F.B. van Duijneveldt, Utrecht). De onderzoekstage werd uitgevoerd binnen het project Structuur-werkingsrelaties van biologische actieve verbindingen (Drs M.L. Verdonk, Prof. dr J. Kroon, Utrecht) bij de vakgroep Kristal- en Structuurchemie van het Bijvoet Centrum voor Bio-moleculair Structuuronderzoek. Daar schreef hij een afstudeerscriptie getiteld Carbohydrate-Protein Interactions in the Crystal Structures of Complexes. Van maart 1993 tot en met mei 1997 was hij assistent-in-opleiding bij de Vakgroep Medicinal Chemistry, Utrecht Institute for Pharmaceutical Sciences, Faculteit Farmacie in Utrecht (Prof. dr J.P.A.E. Tollenaere, Prof. dr J. Kroon (scheikunde)). De resultaten van zijn promotieonderzoek zijn beschreven in het onderhavige proefschrift. In 1994 volgde hij de eerste Guide course on Drug Design and Computational Medicinal Chemistry, Utrecht (Dr E.E. Moret en Prof. dr J.P.A.E. Tollenaere). In 1996 nam hij deel aan de 24th Course of the International School of Crystallography in Erice, Italië, getiteld, Experimental and Computational Approaches to Structure-Based Drug Design. Als onderdeel van zijn promotieonderzoek volgde hij in 1996 een korte stage bij de groep Theoretische Medicinale Chemie (Prof. dr J.P.A.E. Tollenaere) van de Janssen Research Foundation te Beerse, België

LIST OF PUBLICATIONS

M.L. Verdonk, G.J. Boks, H Kooijman, J.A. Kanters and J. Kroon, Stereochemistry of charged nitrogen-aromatic interactions and its involvement in ligand-receptor binding, *J. Comput.-Aided. Mol. Design* , **1993**, *7*, 173-182

I.M. Rigter, M. van Groningen, G.J. Boks, J.P. Tollenaere, K.I. Trollope and J.G. Vinter, Comparison of NK_1 antagonists based on properties of their molecular fields, *Pharmacy World & Science* , **1994**, *16* , Suppl. H15 (abstract)

G.J. Boks, J. Pippel and J.P. Tollenaere, The role of aromatic groups in the NK_1 peptide and non-peptide antagonist pharmacophore, *Pharmacy World & Science*, **1995**, *17*, Suppl. G5 (abstract)

G.J. Boks, J.P. Tollenaere and J. Kroon, Possible ligand-receptor interactions for non-peptide NK_1 antagonists as observed in their crystal structures, *Bioorganic & Medicinal Chemistry*, **1997**, 5, 535-547.

G.J. Boks, D. R. Boer, J.A.W. Kruijtzer, R.M.J. Liskamp, J.P. Tollenaere, J. Kroon en A. Schouten, Synthesis and crystal structure of an aspartame related peptoid (*N*-substituted glycine) peptidomimetic, **1997**, *in preparation.*

G.J. Boks, A.S. Kostense, E.E. Moret, J. Kroon and J.P. Tollenaere, Geometrical features of tertiary and secondary amide bonds in crystal structures: implications for peptoids, **1997**, *in preparation.*

G.J. Boks, A.S. Kostense, E.E. Moret, J. Kroon and J.P. Tollenaere, Putative bioactive conformations for characteristic peptoid fragments, **1997**, *in preparation.*