Theoretical Studies of Cyclic Octapeptides from the Marine Ascidian *Lissoclinum patella*

by

Bruce Forbes Milne

BSc Hons. Aberdeen 1998

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Chemistry
University of Aberdeen

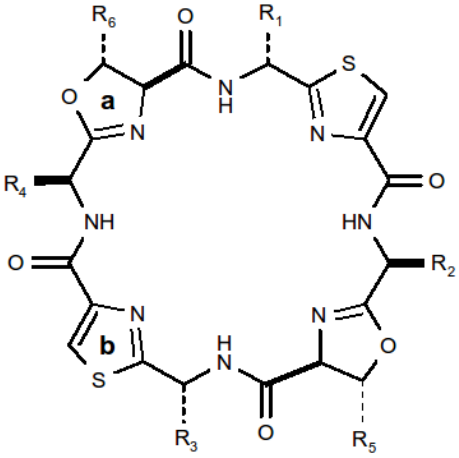
2002

Declaration

This thesis, which is submitted for consideration for the degree of Doctor of Philosophy, is a record of research carried out in the Department of Chemistry, University of Aberdeen, under the supervision of Dr. Marcel Jaspars. It is believed to be original except where due reference has been made and has not been presented for any other higher degree.

Abstract

A comparative theoretical study of oxazole and thiazole was carried out in order to evaluate the performance of several theoretical methods that might be useful in studies on the much larger patellamides (**1**-**8**), a family of octapeptides from the marine ascidian *Lissoclinum patella*. Several density functional theory (DFT) methods were included in this study and it was found that whilst all the functionals tested performed well, HCTH and B3LYP gave results that were better than those obtained at the Hartree-Fock level and equal to or better than the MP2 method but at reduced computational cost. As a result of this preliminary work the B3LYP functional was chosen and the 6-31G(d,p) basis set selected for use in the patellamide study. This level of theory was also chosen for studies of the Cu^{2+} binding site in the patellamides with the SBKJC effective core potential basis.



Patellamides:

1 Ascidiacyclamide R_1=R_3=D-Val R_2=R_4=L-Ile
2 Patellamide A R_1=R_3=D-Val R_2=R_4=L-Ile
3 Patellamide B R_1=D-Phe R_2=L-Leu R_3=D-Ala R_4=L-Ile
4 Patellamide C R_1=D-Phe R_2=L-Val R_3=D-Ala R_4=L-Ile
5 Patellamide D R_1=D-Phe R_2=R_4=L-Ile R_3=D-Ala
6 Patellamide E R_1=D-Phe R_2=R_4=L-Val R_3=D-Ile
7 Patellamide F R_1=D-Phe R_2=R_4=Val R_3=Ile
8 Patellamide G R_1=D-Phe R_2=L-Ile R_3=D-Ala R_4=L-Ile

Note: R_5=R_6=CH_3 for all compounds except patellamide A (**2**) where R_5=CH_3 and R_6=H

Equilibrium geometries and energetics were obtained for a number of conformers/rotamers of ascidiacyclamide (**1**) and patellamides A (**2**), C (**4**) and D (**5**) at the B3LYP/6-31G(d,p) level and were used to help to rationalize the empirical observations that substitutional symmetry appears to be responsible for determining the conformational (folding) preferences in these peptides. The results showed that the energy change on folding ($\Delta E_{folding}$) was positive for all four peptides but that the various substitutions significantly reduced the magnitude of $\Delta E_{folding}$ compared with the C_2 symmetric **1**. Desymmetrization of the oxazoline rings in **2** gives rise to changes in the β-turn forming portions of the macrocycle, leading to improved folding at these positions.

Alternatively, the local structure around the two thiazole rings containing regions can be altered by substitution at R_1 and R_3, leading to improved closure of the macrocucle outwith the β-turn and improved π-stacking between these moieties and shorter hydrogen bonding distances in other areas of the macrocycle.

An investigation into the effects of intra- and intermolecular charge-transfer (CT) on the patellamides was performed using descriptive DFT and the conformational consequences of CT in the patellamides discussed. In **4** and **5**, intermolecular CT is seen to reduce $\Delta E_{folding}$ whereas in the more symmetrical **1** and **2** the $\Delta E_{folding}$ is increased. The dipole moment, also thought to be important in previously suggested intermolecular interactions, was shown to be reduced in magnitude and rotated through 180º on folding of the macrocyclic structure. The possible effect of this on the properties of the patellamides is discussed.

Optimization of the geometry of a model of the Cu^{2+} (TAO) binding site found in the patellamides was used to investigate the preferred coordination environment of the metal with the exogenous ligands H_2O, MeOH, Cl^- and OH^-. The preferred coordination number of the copper was found to be five with all ligands except hydroxide ion, the presence of which lead to a four-coordinate environment. A strong directing effect was observed for negatively charged ligands, moving them into the correct position for form (potentially) reactive bridging ligands between the metal centers in patellamide/dicopper complexes.

The reactivity of the copper center was investigated at the B3LYP/6-31G(d,p) level by evaluating the condensed Fukui functions for different forms of chemical attack. Of the two negatively charged ligands included in the study, chloride was seen to inhibit reactivity whilst hydroxide acted to promote reactivity when compared to the H_2O and MeOH systems (the reactivities of which were almost identical). This final observation seems to go some way to explain the variability found in empirical observations into the reactivity of the patellamide/copper systems.

Acknowledgments

Mainly, I have to thank my supervisor Dr. Marcel Jaspars who gave me the opportunity to study the subjects that were of interest to me, even when the connection between them and the research being performed in the rest of the group seemed tenuous.

My family have been an invaluable source of moral and (sometimes more desperately needed) financial support throughout my all of my studies and deserve a big 'thank you'.

I also want to thank my girlfriend, Manuela, for being so patient as I slowly crept towards the end of my studies. Without her understanding the last few years would have been much more difficult.

Dr. Linda Morris, who worked in Marcel's group before and during the time that I was there, provided a great deal of useful information on the patellamides and the experimental methods used to study them. Without our conversations my grasp of the previous work in this field would have been much worse and the last few years wouldn't have been such a laugh (although the countless cups of coffee we drank during our chats may have left me permanently damaged)

Dr. Martyn F. Guest, who was kind enough to take an interest in my work and find time on the computing facilities at the CLRC Daresbury Laboratories without which a large part of the work in this thesis would have been impossible.

The EPSRC for funding the present work.

I especially want to thank Dr. John H. Binks, lecturer at Aberdeen University during my undergraduate years, who gave me my first real introduction to theoretical chemistry and helped/made me understand some of it despite my appalling grasp of the mathematics involved.

Finally, I would like to thank all the other people who helped in one way or another over

the last few years and made the whole PhD thing that bit easier.

Table of Contents

Part 2
Density Functional Study of Patellamide/Cu^{2+} Complexes

Introduction

Chapter 4
Density Functional Study of Cu^{2+} Binding in the Patellamide TAO Binding Site

Chapter 5
Density Functional Investigation of the Reactivities of TAO/Cu^{2+} Complexes

List of Figures

Abbreviations

A	First electron affinity
Ala	Alanine
ASC	Ascidiacyclamide
B	Becke's approximate GGA exchange functional
CCSD(T)	Coupled-cluster single, double and (non-iterative) triple excitation method
CD	Circular dichroism
$CDCl_3$	Deuterated chloroform
CT	Charge-transfer
Cys	Cysteine
DFT	Density functional theory
DNA	Deoxyribonucleic acid
ED	Electron diffraction spectroscopy
EDTA	Ethylene diamine tetra-acetic acid
EPR	Electron paramagnetic resonance spectroscopy
Et_3N	Triethylamine
GGA	Generalized gradient approximation in DFT
FMO	Frontier molecular orbital (method)
HC	Haemocyanin
HCTH	Approximate GGA exchange-correlation functional of Hamprecht, Cohen, Tozer and Handy
HF	Hartree-Fock
HOMO	Highest occupied molecular orbital
I	First ionization potential
Ile	Isoleucine
KS	Kohn-Sham method in DFT
LDA	Local density approximation in DFT
LUMO	Lowest unoccupied molecular orbital
LYP	Approximate correlation functional of Lee, Yang and Parr
MAD	Mean absolute deviation
MeOH	Methanol
MP2	Møller-Plesset 2^{nd} order perturbation theory
MP4(SDQ)	Møller-Plesset 4^{th} order perturbation theory (including single, double and quadruple excitations)
MD	Molecular dynamics
MM	Molecular mechanics
MS	Mass spectrometry
MSF	Mean scaling factor
MW	Microwave spectroscopy
NMR	Nuclear magnetic resonance
NMR-MD	Molecular dynamics using restraints from experimental NMR data
nOe	Nuclear Overhauser effect
ox	Oxalate
P86	Perdew's approximate GGA correlation functional
PatA	Patellamide A
PatC	Patellamide C

PatD	Patellamide D
Phe	Phenylalanine
phen	Phenanthroline
PO	Phenoloxidase
proPO	Pro-phenoloxidase
q^M	Mulliken partial charge
RHF	Restricted Hartree-Fock molecular orbital method
RKS	Restricted Kohn-Sham DFT method
S	Slater's approximate exchange functional
SCF	Self-consistent field
SD	Standard deviation
Ser	Serine
SF	Scale factor
SMP	Shared memory processing
SQP	Square planar
TAO	Copper binding motif in the patellamides
TBP	Trigonal bipyramidal
Thr	Threonine
TM	Transition metal
UHF	Unrestricted Hartree-Fock molecular orbital method
UKS	Unrestricted Kohn-Sham DFT method
VWN	Approximate LDA correlation functional of Vosko, Wilk and Nusair
Val	Valine
XRD	X-Ray diffraction

Part 1

Introduction

General introduction to ascidian biology

Figure 1. Examples of *Lissoclinum patella* colonies (green masses in center and upper right-hand corner of image) found on surface of coral reef in the Spermonde archipelago, SW Sulawesi, Indonesia. Image downloaded from the Dutch Ascidian Website (www.ascidians.com).

The compounds central to the present work are natural products isolated from the marine organism *Lissoclinum patella* (Figure 1). The phylogenetic profile of *L. patella* is given below:

Phylum	*Chordata*
Subphylum	*Urochordata (Tunicata)*
Class	*Ascidiacea*
Order	*Enterogona*
Family	*Didemnidae*
Genus	*Lissoclinum*
Species	*patella*

Generally, organisms of this type are referred to either as ascidians (from the class name *Ascidiacae*) or tunicates (from the alternative subphylum name *Tunicata*). These

organisms are also commonly known as "sea-squirts" because of their habit of forcefully expelling water when disturbed. Ascidians are filter-feeding animals found throughout the world's oceans and almost exclusively remain fixed to one spot in relatively shallow water throughout their adult lives (juvenile ascidians on the other hand are almost exlusively free-swimming in the early stages of their development – see introduction to Part 2).[1,2] A schematic of the general body plan of an adult ascidian is given in Figure 2. The body of the ascidian is covered in a tough, leathery "tunic" made largely from a material called tunicin (structurally related to cellulose) and the structural features of the body cavities are largely formed from folds in this outer layer. The largest individual feature in the body of the adult ascidian is the pharynx which has a mouth-like opening called the branchial syphon situated (usually) at the top of the body. It is the contraction of the muscles surrounding this pharyngeal sac that produces the jet of water leading to the common name for these organisms. Food particles entering through the branchial syphon are collected by cilia lining the pharynx and passed to the digestive system and waste is expelled through the smaller atrial syphon.

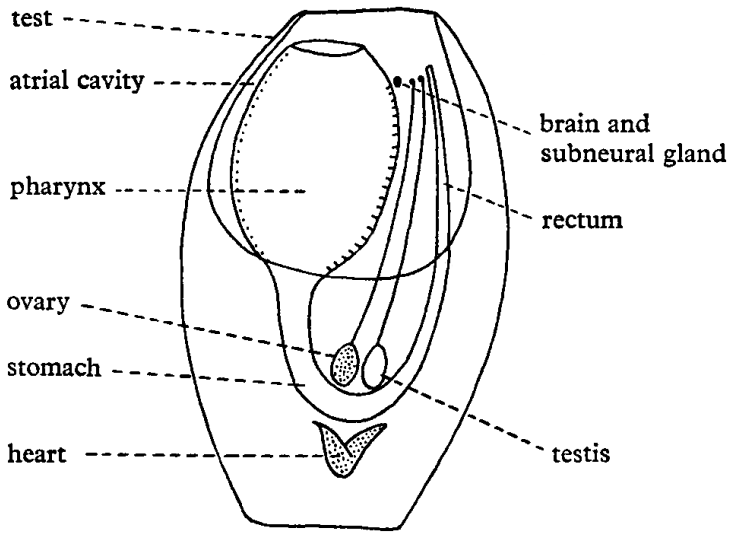


Figure 2. Schematic of the general body plan of an adult ascidian. Modified from Borradaile, 1963, p721.[2]

In colonial species such as *L. patella*, a large number of adults (the term zooid is often used for these colonial individuals) cluster together so that their outer "tunics" merge to form a continuous whole. Commonly in such cases, the organization of the body of each

zooid has become modified so that only the branchial syphon (mouth) is evident on the surface of the colony. Instead of being pointing directly out into the surrounding water, the openings of the atrial syphons of the zooids are gathered together into a common space within the colony from which the waste is ejected. The branchial openings of the *L. patella* zooids and other colonial structures are not easily seen in Figure 1, however an example of another colonial ascidian, *Didemnum molle* (order *Enterogona*, family *Didemnidae*), in which these features can be easily seen is given in Figure 3.

Figure 3. Colonial ascidian *Didemnum molle*. The branchial openings (mouths) of the individual zooids are evident as small dark green spots on the surfaces of the colonies. The central shared atrial cavity is clearly visible through the openings on the tops of the colonies. Image from the Reader's Digest Book of the Great Barrier Reef.[3]

Ascidians display a wide range of colouration, mainly due to pigments within the body of the individual or colony, but the dark green colour seen in the images of *L. patella* and *D. molle* is in fact due to the presence of photosynthetic microorganisms which live in a symbiotic relationship with the host ascidian. This is a common feature of many colonial ascidians with the location of the symbionts varying from species to species. In some

ascidian species the symbiont is found on the surface of the test (tunic) or inside the body cavities (*D. molle* is a good example of the latter, with the dark green colour of the symbiont evident in the common atrial openings – Figure 3) whilst in other ascidians the symbiont is actually found within the tissues of the host. This variation may reflect the varying degree to which the symbiotic relationship has developed in the different species of ascidians (in some species the "symbiont" can actually be removed from the surface of the ascidian by gently brushing the colony).[3]

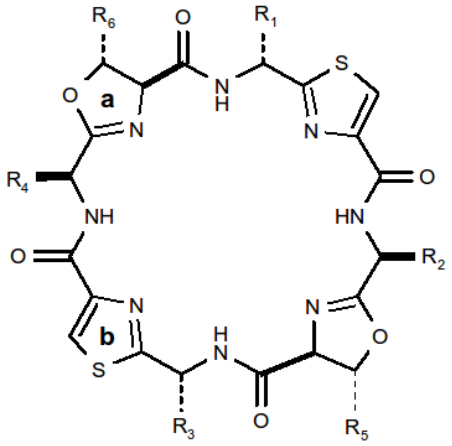
In *L. patella* the symbiont is a single-celled blue-green algae of the genus *Prochloron* which is localized within the host tissue surrounding the zooids. The *L. patella* zooids themselves are located near the surface of the colony and careful dissection of a colonial mass by Salomon and Faulkner, 2002, showed that although *Prochloron* is found living within the tunic itself, no algal cells are found in the deeper regions of the colony (probably because the penetration of light required for photosynthesis is poor at these deeper levels).[4] A more general localization of *Prochloron* within the *L. patella* colony is evident from Figure 1 where the dark green (*Prochloron*-containing) regions correspond to depressions in the surface of the colony. This localization is again probably related to the conditions required by the algae.

For some time the presence of *Prochloron* has been a complicating factor in the study of natural products obtained from *L. patella* as it is unclear whether any compounds isolated from these colonies came from the ascidian or its algal symbiont. Degnan *et al.*, 1989, found that a number of compounds isolated from intact *L. patella* colonial masses could be obtained in equal or greater amounts from the *Prochloron* cells if these were first removed from the host.[5] A similar result was obtained by the authors in a study of the related *Lissoclinum bistratum* which also contains *Prochloron*.[6] In contrast to these results the study of Salomon and Faulkner, 2002, in which careful dissection of the colony followed by centrifugal separation of any algal cells and careful washing to remove any *L. patella* debris remaining showed that the compounds described by Degnan *et al.* could be obtained from all areas of the *Prochloron*-free colonial mass but not from the algal cells themselves.[4] Whilst this last result cannot rule out completely the possibility that the symbiont is producing the secondary metabolites and that these are then transported to the

host, the degree of care taken during the separation seems to suggest that if transport of the compounds between the symbiont and host is occurring then the process in one-way and the rate is so fast that the compounds only exist in the algal cells in undetectable amounts.

The patellamides: cyclic peptide-based natural products isolated from *L. patella*

The patellamides (**1-8**, Figure 4) are a family of biologically active cyclic octapeptides with [24]azacrown-8 structure isolated from *L. patella* and are characterized primarily by the presence of two thiazole and two oxazoline rings fused with the peptide backbone.



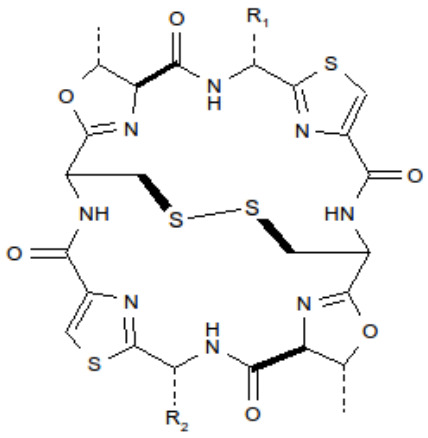
Patellamides:

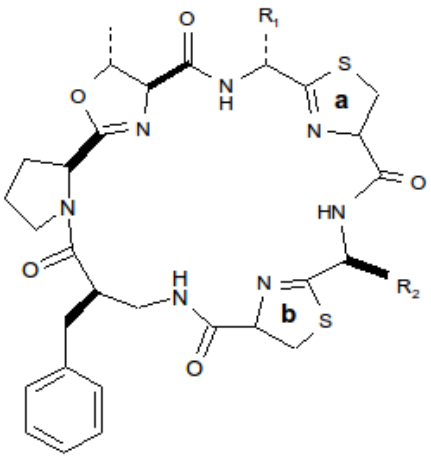
1 Ascidiacyclamide R_1=R_3=D-Val R_2=R_4=L-Ile
2 Patellamide A R_1=R_3=D-Val R_2=R_4=L-Ile
3 Patellamide B R_1=D-Phe R_2=L-Leu R_3=D-Ala R_4=L-Ile
4 Patellamide C R_1=D-Phe R_2=L-Val R_3=D-Ala R_4=L-Ile
5 Patellamide D R_1=D-Phe R_2=R_4=L-Ile R_3=D-Ala
6 Patellamide E R_1=D-Phe R_2=R_4=L-Val R_3=D-Ile
7 Patellamide F R_1=D-Phe R_2=R_4=Val R_3=Ile
8 Patellamide G R_1=D-Phe R_2=L-Ile R_3=D-Ala R_4=L-Ile

Note: R_5=R_6=CH_3 for all compounds except patellamide A (**2**) where R_5=CH_3 and R_6=H

Figure 4. Schematic diagram of the molecular structure of the patellamides (**1-8**). Five-membered ring systems: **a**=oxazoline, **b**=thiazole.

Similar features are observed in several other families of cyclic peptides isolated from *L. patella* including the ulithiacyclamides[7-9] (**9** & **10**, Figure 5) and the lissoclinamides[5,10,11] (**11-20**, Figure 5). These 5-membered heterocycles are formed by the condensation of cysteine and serine/threonine residues, respectively, with neighbouring residues within the macrocyclic ring (Figure 6).[12]





Ulithiacyclamides:

9 Ulithiacyclamide A R_1=R_2=D-Leu
10 Ulithiacyclamide B R_1=D-Ile R_2=D-Phe

Lissoclinamides:

11 Lissoclinamide 1 R_1=L-Val R_2=D-Ile **a**=**b**=thiazole
12 Lissoclinamide 2 R_1=D-Ile R_2=D-Ala **a**=thiazoline **b**=thiazole
13 Lissoclinamide 3 R_1=D-Ile R_2=D-Ala **a**=thiazoline **b**=thiazole
14 Lissoclinamide 4 R_1=L-Val R_2=D-Phe **a**=thiazoline **b**=thiazole
15 Lissoclinamide 5 R_1=L-Val R_2=D-Phe **a**=**b**=thiazole
16 Lissoclinamide 6 R_1=D-Val R_2=D-Phe **a**=thiazoline **b**=thiazole
17 Lissoclinamide 7 R_1=D-Val R_2=D-Phe **a**=**b**=thiazoline
18 Lissoclinamide 8 R_1=Val R_2=Phe **a**=thiazoline **b**=thiazole
19 Lissoclinamide 9 R_1=L-Ile R_2=D-Val **a**=thiazoline **b**=thiazole
20 Lissoclinamide 10 R_1=L-Ile R_2=L-Ile **a**=thiazoline **b**=thiazoline

Figure 5. Schematic diagram of the molecular structure of the ulithiacyclamides (**9** & **10**) and the lissoclinamides (**11**-**20**).

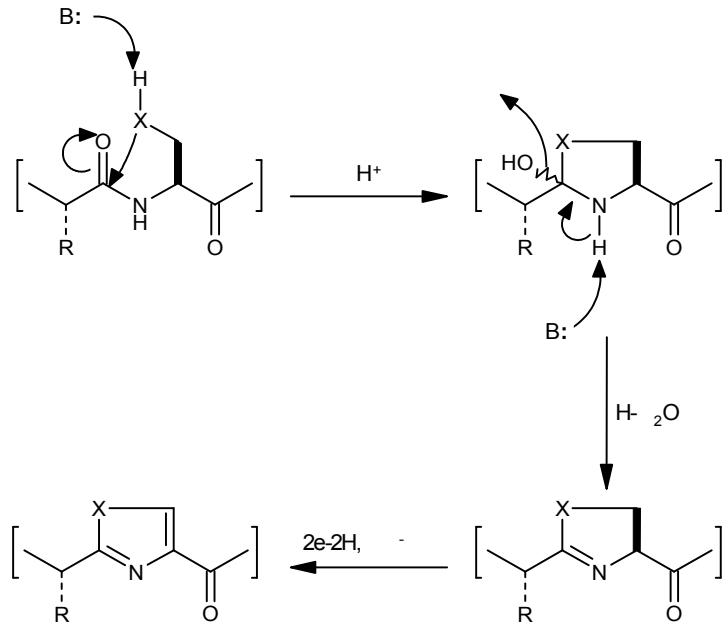


Figure 6. Suggested biosynthetic mechanism for formation of oxazoline (X=O)/thiazoline (X=S) by

cyclodehydration of serine/cysteine followed by $2e^-$ reduction to form oxazole/thiazole.[12]

The condensed residues are incorporated into the peptide in an alternating fashion and are separated from each other by 'spacer' groups, each of which consist of a single D- or L-amino acid residue in which the side chains remain unmodified. All of the patellamides have these spacer residues arranged in a D-L-D-L sequence and this also contributes significantly to the nature of the observed conformations. It is the type and distribution of these residues that distinguishes the different members of the family. The heterocyclic components of the patellamide structures convey structural rigidity and place significant constraints on the range of dynamical motion displayed by these peptides.

Conformational preferences and properties of the patellamides

Previous studies of the patellamides have shown that these compounds universally populate either of two possible overall conformations (Figure 7). The conformational preferences observed for different patellamides in a variety of environments are summarized in Table 1 along with the experimental techniques that have been applied to these compounds. As well as solid state structures obtained by X-ray diffraction (XRD)[13,14] a number of workers have employed techniques such as nuclear magnetic resonance (NMR)[13,15-18] and circular dichroism (CD)[17-19] to investigate the conformational preferences of the patellamides in a variety of solvents.

The varying preferences displayed by the patellamides for either the open "square" form or the closed "figure-of-eight" form seem to be related to the degree and symmetry of substitution of the unmodified amino acid residues within the macrocycle. The solution environment of the solvated patellamides also appears to contribute significantly to whether the open or closed conformation is observed, with polar solvents seeming to promote population of the open conformer and non-polar solvents promoting population of the closed conformer (see Table 1).

The empirically-derived rationale outlined above provides a very useful insight into the origins of the conformational preferences displayed by the patellamides. It has already lead to a number of investigations involving synthetic analogues of the patellamides in which it has been possible to vary the substitution pattern of the unmodified residues

and/or neglect the cyclization of the serine/threonine residues to form the oxazoline ring units.[16,18] Perhaps the most significant result of these studies has been to show that the oxazoline rings are vital in the control of the folding properties of the patellamides, with all compounds lacking these features exclusively adopting the closed conformation (irrespective of the nature of the unmodified residues or the degree of symmetry imposed by their inclusion).

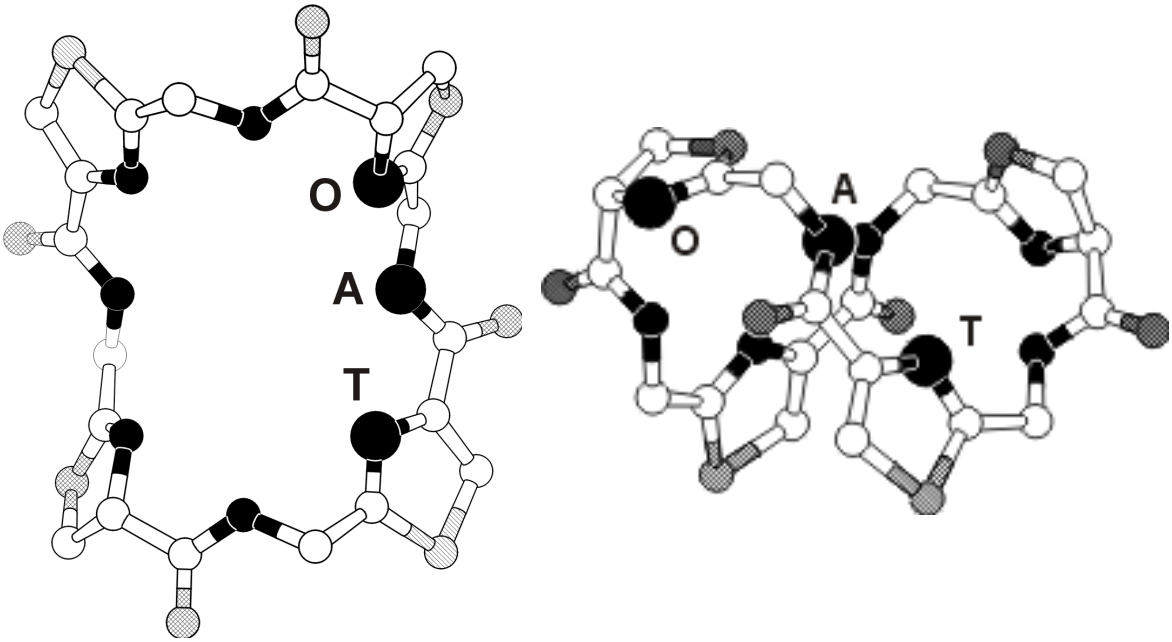


Figure 7. Open (left) and closed (right) conformations adopted by the patellamides. Letters correspond to the nitrogen atoms comprising the copper binding site in these peptides (see Chapter 4). T=thiazole, A=amide and O=oxazoline. Figures reproduced from Morris *et al.*, 2002.[20]

Compound	Conformation	Method	Solvent	Reference
1 Ascidiacyclamide	open	Crystal	C_6H_6	13
	open	NMR	$CDCl_3$	13
2 Patellamide A	open	Crystal	MeOH	14
	open	CD (299K)	MeOH	19
	closed	CD (199K)	MeOH	19
	open	CD	MeOH	17
	closed	NMR-MD	$CDCl_3$	17
3 Patellamide B	closed	NMR-MD	$CDCl_3$	15
	closed	CD	MeOH	19
4 Patellamide C	closed	NMR-MD	$CDCl_3$	15
	closed	CD	MeOH	17
5 Patellamide D	closed	Crystal, MD	MeOH	25
6 Patellamide E	closed	CD	MeOH	19

Table 1. Experimentally observed conformations for natural patellamides (**1**-**6**).

The study of the patellamides has been put on a semi-quantitative footing through the use of molecular mechanics/dynamics (MM/MD) calculations. nOe-restrained MD simulations have been used by several groups in the interpretation of NMR data obtained for a number of patellamides in various solvents.[15-17] Unrestrained MD simulations have also provided information on the vital role of the oxazoline units in the folding process as well as the mediating effects of solvent polarity (see Appendix A; Milne *et al.*, 2002).[17] Close study of the unrestrained MD trajectories has lead to the proposal of a four step mechanism for the transition between the open and closed conformations (Figure 8).[17]

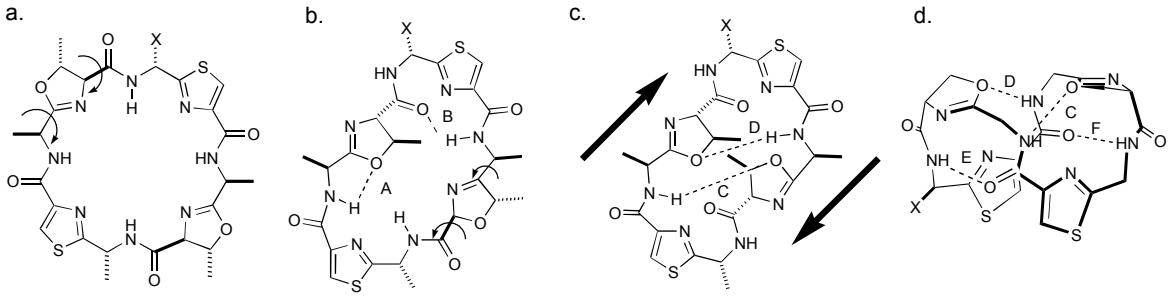


Figure 8. Folding mechanism in the patellamides observed in unrestrained MD simulations.[17] Side groups have been removed to increase clarity and an X added to aid in following the folding process. Initial rotation of one oxazoline ring in the open conformation (a) leads to formation of hydrogen bonding interactions A and B (b). This is followed by rotation of the second oxazoline leading to two more hydrogen bonding interactions , C and D (c). Subsequently, the whole structure undergoes a shearing motion (c) which allows the patellamide to assume the closed conformation (d).

These conformational properties are of interest from a biological/medicinal perspective as investigations into the bio-activity of the different patellamides have highlighted their cytotoxicity and in-vivo anti-tumor activity, the strengths of which vary within the series.[5,21] Patellamide D (**5**),[5] which is found in the closed conformation, has been shown to enhance the effects of antibacterial compounds against multi-drug-resistant bacteria in-vitro.[22] This property has also been demonstrated for two other members of the patellamide family that are known to favour the closed conformation , patellamides B (**3**) and C (**4**).[23]

In addition to the biological properties mentioned above, the physical properties of the patellamides have attracted attention from researchers. The interaction of ascidiacyclamide (**1**) with model phospholipid in a buffered solution was investigated using differential scanning calorimetry and it was found that on addition of **1** to this system alterations in the

phase-transition behaviour of the phospholipid occurred suggesting that the peptide was having some effect on the phospholipid structures.[13] It has also been suggested that the patellamides may have electron-transfer rôles in-vivo[24] and that charge-transfer between the thiazole ring units may be important in the stabilization of the closed conformation of these peptides[25]. The metal binding properties of the patellamides has also been the subject of a great deal of work, with unusually high selectivity for Cu^{2+} over other divalent metal cations being observed (this is the subject of Part 2 of this thesis and a review of previous experimental and computational work is given in the general introduction to Part 2 and at the beginning of Chapter 4).

Present work: computational/theoretical investigation of conformational properties of the patellamides

The detailed interpretation of the factors influencing the conformational preferences displayed by the patellamides (**1**-**8**) using physical methods has been hampered to some degree by the variability observed in the results of these studies. The most reliable geometrical data obtained for these compounds has come from X-ray diffraction (XRD) studies which have provided considerable insight into their (thermally averaged) equilibrium structures.[14, 26-29]

Unfortunately, rigorous comparisons of the patellamides that have been subjected to this type of study are complicated by the fact that, even for two molecules within a single unit cell, the forces acting on the structures due to the presence of the neighbouring molecules lead to variable distortions that become evident if comparisons of bond lengths, angles or dihedrals are attempted. Similarly, the structures obtained previously using nOe-restrained molecular dynamics (nOe-MD) techniques are subject to variations in the nature and degree of solvation of the molecules and a large amount of time-averaging of the flexible patellamide structure during the collection of quantities of data sufficient to allow reliable assignment of the signals observed. The details of this data are then further obscured by the use of empirical force fields in the modelling of the structures and by the statistical treatment of the results of these calculations.

Highly accurate structural data that are not subject to the environmental effects mentioned

above may be obtained for small, rigid molecules through the use of gas phase spectroscopic techniques such as microwave (MW) or electron diffraction (ED) spectroscopy. However, this approach is not particularly well-suited for studying the patellamides due to their large size and the flexible nature of their structures. Alternative procedures that are, in principle at least, capable of producing structural information of comparable precision to the MW and ED techniques come from the field of computational theoretical chemistry and with the recent developments in the performance of computers these methods are now becoming applicable even to large biomolecular structures.

As mentioned previously, some attempt has been made through the use of force-field calculations to quantify the differences between the members if the patellamide family. This approach suffers from the significant drawback that the potentials used are themselves empirical with the result that if a force-field has not been calibrated specifically for the molecule of interest, then any results obtained through its use must be treated with caution. For this reason it becomes difficult to evaluate the importance of different contributions (e.g., solvation energy, changes in steric energy due to amino acid substitutions) to the thermodynamic stability of the different conformations.

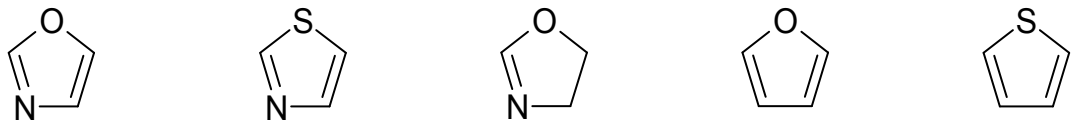
In the present work it was decided to study the conformational trend in a subset of the natural patellamide family using a first principles computational model to study gas phase energetics and geometries in the absence of the complicating factors associated with force-field modelling. The software chosen to perform this work was the GAMESS-UK package available from the CLRC Daresbury Laboratory.[30] Version 6.3 of the package was provided by Drs. Martyn F. Guest and Paul Sherwood at Daresbury and installed on a dual 800MHz Pentium III system that had been obtained for this purpose (earlier, a Silicon Graphics O2 workstation was used but the processing power was insufficient for the studies planned at Aberdeen). The RedHat v7.0 linux operating system was used in order to make full use of the computational options available due to the shared memory processing (SMP) capabilities of this machine. All preliminary work for the patellamide study (and all work on the Cu^{2+} complexes in Part 2) was carried out using this system.

Chapter 1

Comparative Theoretical Study of Oxazole and Thiazole

Introduction

In order to gain some experience with GAMESS-UK and to assess the performance of theoretical methods that might be applicable to the conformational study on the full patellamide structures it was decided to carry out a comparative study on the five-membered heterocycles oxazole (**21**) and thiazole (**22**). The rationale behind these choices was that both molecules were small enough to allow the study to be carried out on the dual-processor workstation and both were similar to structural subunits of the patellamides. In addition, accurate experimental data were available from microwave spectroscopic (MW) studies on both molecules.[31-33] Oxazoline (**23**) would have been of more direct relevance than oxazole but no experimental data of sufficient quality could be found in the literature for this compound meaning that assessment of the results of theoretical calculations would have been more difficult.



12 **22** **32** **42** **52**

It was also found that oxazole and thiazole, along with their parent heterocycles furan (**24**) and thiophene (**25**), had been previously investigated using theoretical methods.[31,33-37] The wavefunction-based methods used in the studies on **21** and **22** included the Hartree-Fock self consistent field (HF-SCF) method[33,35-37] and post-HF Møller-Plesset perturbation theory at second- (MP2)[31,33,35-37] and fourth-order (MP4(SDQ))[36] levels (the MP4(SDQ) correction was only used for the calculation of properties at the MP2 geometries and not for the geometry optimizations themselves). Density functional theory (DFT) had been applied to oxazole (**21**) in the form of the hybrid HF-DFT functional B3LYP.[38] The most extensive DFT study was that of El-Azhary and Suter, 1996,[37] on compounds **24** and **25** which included the SVWN functional (local density approximation (LDA)), BLYP (generalized gradient approximation (GGA)) and again the hybrid B3LYP functional. The BVWN functional (GGA exchange with LDA correlation) was also used in this study.

El-Azhary and Suter, 1996, also investigated the effects of basis set quality on the results obtained with the different methods.[37] The authors found that increasing the size of the

basis set from 6-31G(d,p) to cc-pVDZ actually lead to a worse geometrical description of **24** and **25**. An example of this is the C=C double bond in furan (**24**) where deviations of +0.003 Å (MP2) and 0.000 Å (B3LYP) were found with the 6-31G(d,p) basis. On changing to the larger cc-pVDZ set these deviations became +0.014 Å (MP2) and +0.003 Å (B3LYP). The authors also used the even larger cc-pVTZ set, however they were unable to complete the MP2 calculation with this basis (presumably because of excessive time requirements). The B3LYP/cc-pVTZ combination showed even greater deviation than B3LYP/cc-pVDZ with the error in the C=C length becoming -0.007 Å. The deviations found for the equivalent bond in thiophene (**25**) when using these basis sets were almost identical. In the heterovalent C-O/S bond lengths the B3LYP results showed significant improvement on increasing the basis set size but for the MP2 method only a 0.001 Å improvement was observed for the C-O length with the C-S length becoming 0.011 Å too large on going to the cc-pVDZ basis (as opposed to the MP2/6-31G(d,p) deviation of 0.000 Å).

In general, these studies found that in the calculation of equilibrium geometries the performance (measured against the experimental data) of the HF method was inferior to both the MP2 and DFT correlated levels of theory. Of the correlated methods, MP2 and B3LYP gave similar results with the DFT method providing the best average fit to the experimental data. The performance with respect to vibrational frequencies was compared in terms of the scaling factor (SF) required to bring the calculated values into line with those obtained experimentally. DFT again provided the most accurate results with the scaling factor for the B3LYP results being closer to unity than the factors required for either the HF or MP2 methods. In the calculations on **24** and **25** these scale factors were 0.819 (HF/6-31G(d,p)), 0.891 (MP2/6-31G(d,p)) and 0.928 (B3LYP/6-31G(d,p)).[37]

Within the context of the present work the quality of the DFT results was of particular interest as the favourable scaling properties of DFT methods in general ($\sim M^4$, where M is the number of basis functions used)[39] meant that this approach was well suited to studies on large systems like the patellamides. The uncorrelated HF method scales formally as M^4 (depending on the exact nature of the calculation this can drop to M^1).[40] The least demanding method for the introduction of electron correlation effects as a post-HF

correction is MP2 which scales as M^5.[40] Whilst the difference in scaling between the DFT and MP2 methods may not represent a huge saving in studies on small molecules with moderately sized basis sets, performing correlated electronic structure calculations on medium to large systems rapidly becomes too computationally expensive even when minimal basis sets are used.

Methods

As DFT had only been applied to oxazole (**21**) and thiazole (**22**) previously using the B3LYP functional[41] it was decided to perform a comparative study similar to that carried out on furan (**24**) and thiophene (**25**) by El-Azhary and Suter.[37] A number of approximate density functionals available in GAMESS-UK were selected to represent the LDA, GGA and hybrid HF-DFT levels of theory. Several commonly used functionals were chosen: SVWN for the LDA approximation,[42] BP86 [43,44] and BLYP [43,45] for GGA and the hybrid HF-DFT functional B3LYP. The less well reported HCTH GGA functional of Hamprecht *et al*., 1998,[46] was also included as a benchmark study on a number of sulfur-containing compounds[47] had shown this functional to provide geometrical and vibrational values superior to those obtained with the more commonly used BLYP GGA functional and the MP2 (frozen core) perturbation method. In the present work HF and frozen core MP2 calculations were also performed for comparison with the DFT results. Several standard basis sets of double- and triple-ζ quality were used in combination with the DFT and wavefunction based methods (Table 2).

Basis set	Valence space quality	Number of Contracted Basis Functions	
		Oxazole	Thiazole
6-31G(d)	Double-ζ	81	85
6-31G(d,p)	Double-ζ	90	94
DZP	Double-ζ	95	103
cc-pVDZ	Double-ζ	90	94
SBKJC(d)	Double-ζ	76	76
SBKJC(d,p)	Double-ζ	85	85
6-311G(d)	Triple-ζ	104	112
6-311G(d,p)	Triple-ζ	113	121
TZP	Triple-ζ	118	125
cc-pVTZ	Triple-ζ	220	224

Table 2. Basis sets used in the study on oxazole (**21**) and thiazole (**22**). The SBKJC basis uses a relativistic effective core potential (ECP) to represent core electrons (*e.g.*, the 1s electrons in carbon) whereas the other basis sets used here contain functions for the explicit description of this region. SBKJC provides a -31G split valence description of main group atoms.[48] **Note:** The fact that the O^N scaling factors for the different methods given in the text are formal rather than operational values is highlighted by the case of calculations with the cc-pVDZ basis which required approximately twice as much time as the same calculation using 6-31G(d,p) despite using the same number of contracted basis functions.

Each combination of theoretical method and basis set was used to calculate the equilibrium geometries of the two test molecules and the results compared with the MW geometries from the literature.[33] Dipole moments were calculated with each method/basis set combination at the corresponding equilibrium geometries and compared with experimental data available in the literature.[49] The 6-31G(d,p) basis was used in combination with each of the theoretical methods to calculate the vibrational frequencies of both test molecules. The experimental frequencies reported in El-Azhary, 1995, were used for evaluation of these results.[33]

Results

Geometries

Due to the large quantity of data obtained, explicit comparisons of individual geometric coordinates are neglected in this section. Instead, the mean absolute deviation (MAD) from the experimental data has been used to provide a summary of these comparisons and to give a simple measure of the performance of the various methodologies. Values relating to heavy atom bond lengths and angles are shown in Tables 3-6.

No data is given for the MP2/cc-pVTZ combination as these calculations did not run properly. This seems to have been due either to a problem with the software or an incompatibility between the method and the basis set as leaving the calculation to run for up to a week made no difference (the initial SCF procedure converged but the subsequent MP2 orbital transformation appeared to lock up and no further progress was observed).

Oxazole

In the calculations on oxazole (**21**) the MAD values relating to bond lengths ranged from 0.001 Å with the B3LYP/6-311G(d) method to 0.029 Å obtained with the BLYP functional and either the SBKJC(d) or SBKJC(d,p) effective core potential basis sets (Table 3). Interestingly, the 6-31G(d) and 6-31G(d,p) basis sets gave some of the best results with all of the correlated methods and increasing the size and/or valence space quality of the basis had very little effect. The relatively poor performance of the cc-pVDZ basis noted by El-Azhary and Suter, 1996,[37] was evident in the results obtained here. With the exception of the RHF method, the values obtained with this basis showed a reduction in accuracy compared with those found with 6-31G(d,p) and in addition required approximately twice as much CPU time. The DZP basis also displayed this tendency. When using the SBKJC-based sets the RHF calculations proved to be the most accurate, closely followed by HCTH and B3LYP. The MAD values for the remaining methods in combination with these ECP basis sets fell in the order SVWN ≈ MP2 < BP86 < BLYP. B3LYP gave the best performance of all the methods studied when combined with the triple-ζ basis sets. HCTH was only slightly less accurate than B3LYP with these quality of basis set. The other methods showed deviations in the order MP2 < SVWN ≈ BP86 < BLYP < RHF.

Regardless of the quality of the basis set used the B3LYP and HCTH density functionals consistently provided more accurate bond lengths than the MP2 method. The worst performances overall were those of the RHF and BLYP methods with SVWN and BP86 falling in the middle of the range.

Basis	RHF	MP2	SVWN	BP86	BLYP	HCTH	B3LYP
Double-ζ							
6-31G(d)	0.019	0.006	0.009	0.011	0.015	0.005	0.002
6-31G(d,p)	0.019	0.006	0.011	0.011	0.014	0.005	0.002
DZP	0.016	0.009	0.011	0.015	0.020	0.007	0.005
cc-pVDZ	0.017	0.009	0.010	0.012	0.016	0.006	0.003
SBKJC(d)	0.009	0.020	0.019	0.023	0.029	0.012	0.015
SBKJC(d,p)	0.009	0.019	0.019	0.023	0.029	0.012	0.015
Triple-ζ							
6-311G(d)	0.020	0.007	0.009	0.009	0.014	0.004	**0.001**
6-311G(d,p)	0.020	0.008	0.009	0.009	0.014	0.004	0.002
TZP	0.020	0.005	0.009	0.009	0.014	0.004	0.002
cc-pVTZ	0.021	n/a	0.010	0.007	0.011	0.004	0.003

Table 3. Oxazole (**21**) heavy atom bond lengths: mean absolute deviations with respect to MW data.[33] Lowest MAD shown in bold type. Values in Å.

Basis	RHF	MP2	SVWN	BP86	BLYP	HCTH	B3LYP
Double-ζ							
6-31G(d)	0.31	0.18	0.27	0.19	0.19	0.20	0.16
6-31G(d,p)	0.30	0.17	0.29	0.20	0.20	0.18	0.16
DZP	0.32	0.21	0.42	0.14	0.21	0.16	0.20
cc-pVDZ	0.36	0.28	0.27	0.13	0.17	0.25	0.16
SBKJC(d)	0.46	0.23	0.21	0.18	0.19	0.24	0.27
SBKJC(d,p)	0.47	0.30	0.22	0.18	0.20	0.24	0.30
Triple-ζ							
6-311G(d)	0.34	0.18	0.26	0.13	0.21	0.18	0.13
6-311G(d,p)	0.34	0.19	0.30	0.14	0.22	0.13	0.17
TZP	0.39	**0.12**	0.38	0.24	0.33	0.14	0.29
cc-pVTZ	0.31	n/a	0.40	0.15	0.26	0.07	0.22

Table 4. Oxazole (**21**) heavy atom bond angles: mean absolute deviations with respect to MW data.[33] Lowest MAD shown in bold type. Values in degrees.

The trend in the calculated heavy atom bond angles for oxazole (**21**; Table 4) was slightly different from that seen in the bond lengths but again B3LYP generally provided better results than the MP2 method. The only exception to this was observed when using the TZP basis which yielded a MAD for heavy atom bond angles of 0.12º with MP2 and 0.29º with B3LYP. The superiority of HCTH over MP2 shown in the bond lengths was slightly lessened with the two methods providing similar results. No significant difference was observed between these methods and the results of the BP86 and BLYP calculations. SVWN and RHF both performed relatively poorly with the DFT method providing a slight

improvement over RHF.

The effects of basis set variations were similar to those seen for the bond lengths. In the DFT calculations no large improvements were seen on increasing the basis set above the polarized 6-31G sets. Similarly, only small improvements were seen for RHF and MP2 when larger basis sets were used. The SBKJC sets gave much more competitive results for bond angles than those observed for bond lengths. The cc-pVDZ basis gave good results with the BP86 functional (MAD = 0.13º but its performance with the other methods was no better than the 6-31G sets.

Thiazole

The MAD values for thiazole (**22**) bond lengths (Table 5) were noticeably higher than those seen for oxazole. The MP2/6-31G(d) combination gave the lowest MAD value of all the methods tested (0.004Å). The overall performance of MP2 made it the most accurate method with SVWN and HCTH showing comparable, if slightly lower, accuracy. Of the DFT methods tested the SVWN functional provided the best performance, closely followed by HCTH and B3LYP. The worst overall performance of the DFT methods was again given by BLYP with the BP86 results lying approximately half-way between the BLYP and B3LYP values. The RHF results were seen to be reasonably competitive when the all-electron double-ζ basis sets were used but deteriorated slightly when the basis quality was increased to triple-ζ. As seen in the calculations on oxazole, the RHF method gave the most accurate results of all the methods when the SBKJC ECP basis sets were employed.

The 6-31G – based basis sets gave the lowest deviations in the double-ζ group when used with the correlated MP2 and DFT methods, with the other sets in this group showing slightly increased deviations from experiment (B3LYP/DZP was the one exception to this rule as it gave the same deviation (0.012Å) as B3LYP/6-31G(d) and B3LYP/6-31G(d,p)). In general, no significant improvement was observed when triple-ζ rather than double-ζ all electron basis sets were used, regardless of the level of sophistication of the theoretical method.

Basis	RHF	MP2	SVWN	BP86	BLYP	HCTH	B3LYP
Double-ζ							
6-31G(d)	0.016	**0.004**	0.006	0.018	0.024	0.008	0.012
6-31G(d,p)	0.016	0.005	0.006	0.018	0.024	0.008	0.012
DZP	0.014	0.010	0.007	0.021	0.027	0.011	0.012
cc-pVDZ	0.017	0.010	0.008	0.020	0.026	0.010	0.013
SBKJC(d)	0.011	0.020	0.025	0.030	0.036	0.016	0.022
SBKJC(d,p)	0.010	0.019	0.025	0.030	0.036	0.016	0.022
Triple-ζ							
6-311G(d)	0.018	0.007	0.008	0.017	0.024	0.010	0.014
6-311G(d,p)	0.018	0.006	0.008	0.017	0.024	0.009	0.014
TZP	0.019	0.006	0.008	0.017	0.023	0.009	0.013
cc-pVTZ	0.014	n/a	0.007	0.011	0.016	0.005	0.008

Table 5. Thiazole (**22**) heavy atom bond lengths: mean absolute deviations with respect to MW data.[33] Lowest MAD shown in bold type. Values in Å.

In the reproduction of the bond angles in thiazole the lowest MAD was obtained with the HCTH/cc-pVTZ combination (0.16º). The next lowest value was that of SVWN/DZP (0.17º). MP2 and BP86 both provided values of 0.20º with this basis set as did the B3LYP functional in combination with the SBKJC(d,p) basis. Surprisingly, the results of the BLYP, HCTH and B3LYP calculations with both the SBKJC sets were the most accurate of all the calculations using these functionals. BLYP again consistently gave the largest MAD values with the all-electron basis sets. This was most obvious when the BLYP/triple-ζ results were compared with the equivalent HCTH and MP2 values (see Table 6). The BLYP values with these basis sets were approximately twice as large as those obtained with the two other methods.

With the exception of the HCTH/cc-pVTZ result mentioned above, the use of triple-ζ basis sets again gave no significant improvement over the double-ζ results. Unlike the previous results for oxazole and for the thiazole bond lengths, the SBKJC basis produced some of the lowest MAD values of all the basis sets used. This was particularly noticeable in the results of the density functional calculations using the HCTH and B3LYP functionals. The cc-pVDZ basis was again found to produce results of lower quality than might be expected when compared with the other double-ζ sets.

Basis	RHF	MP2	SVWN	BP86	BLYP	HCTH	B3LYP
Double-ζ							
6-31G(d)	0.38	0.29	0.26	0.32	0.45	0.29	0.34
6-31G(d,p)	0.37	0.30	0.26	0.30	0.44	0.27	0.34
DZP	0.35	0.20	0.17	0.20	0.32	0.24	0.25
cc-pVDZ	0.45	0.41	0.44	0.32	0.45	0.34	0.36
SBKJC(d)	0.32	0.36	0.26	0.28	0.28	0.23	0.21
SBKJC(d,p)	0.30	0.42	0.24	0.26	0.26	0.22	0.20
Triple-ζ							
6-311G(d)	0.47	0.32	0.37	0.42	0.60	0.33	0.42
6-311G(d,p)	0.45	0.36	0.38	0.44	0.61	0.33	0.40
TZP	0.45	0.35	0.37	0.41	0.56	0.33	0.41
cc-pVTZ	0.35	n/a	0.24	0.26	0.42	**0.16**	0.27

Table 6. Thiazole (**22**) heavy atom bond angles: mean absolute deviations with respect to MW data.[33] Lowest MAD shown in bold type. Values in degrees.

Properties

Dipole moments

The calculated dipole moments for compounds **21** and **22** along with their deviations from the literature values (oxazole (**21**)=1.50 ± 0.03 Debye and thiazole (**22**)=1.61 ± 0.03 Debye)[49] are shown in Tables 7 and 8, respectively. As with the results of the geometry optimizations no value is given for MP2/cc-pVTZ as these calculations were aborted.

The results of the dipole moment calculations on oxazole (**21**) show a strong trend favouring the GGA and hybrid HF-DFT methods over either the LDA or wavefunction-based calculations. In contrast to the results of the geometry optimizations, BLYP was seen to give by far the best overall performance whilst none of the MP2 values fell within the experimental range of 1.50 ±0.03 Debye. HCTH, B3LYP and BP86 gave very similar performances to one another with SVWN showing the worst overall agreement with experiment of the DFT methods tested (although the one result (SVWN/cc-pVDZ) that did fall within the experimental range showed a deviation of 0.00 Debye). The RHF results were similar to those obtained with SVWN and also showed a deviation of 0.00 Debye when the cc-pVDZ basis was used.

Basis	RHF	MP2	SVWN	BP86	BLYP	HCTH	B3LYP
Double-ζ							
6-31G(d)	1.59 (0.09)	1.64 (0.14)	1.58 (0.08)	**1.53 (0.03)**	**1.48 (-0.02)**	**1.53 (0.03)**	**1.53 (0.03)**
6-31G(d,p)	1.58 (0.08)	1.62 (0.12)	1.57 (0.07)	**1.52 (0.02)**	**1.48 (-0.02)**	**1.52 (0.02)**	**1.53 (0.03)**
DZP	1.59 (0.09)	1.65 (0.15)	1.60 (0.10)	1.54 (0.04)	**1.50 (0.00)**	1.54 (0.04)	1.55 (0.05)
cc-pVDZ	**1.50 (0.00)**	1.57 (0.07)	**1.50 (0.00)**	1.44 (-0.06)	1.40 (-0.10)	1.45 (-0.05)	1.45 (-0.05)
SBKJC(d)	1.63 (0.13)	1.70 (0.20)	1.61 (0.11)	1.57 (0.07)	**1.51 (0.01)**	1.60 (0.10)	1.56 (0.06)
SBKJC(d,p)	1.62 (0.62)	1.68 (0.18)	1.60 (0.10)	1.56 (0.06)	**1.51 (0.01)**	1.59 (0.09)	1.56 (0.06)
Triple-ζ							
6-311G(d)	1.56 (0.06)	1.64 (0.14)	1.57 (0.07)	1.17 (-0.33)	**1.48 (-0.02)**	**1.51 (0.01)**	**1.52 (0.02)**
6-311G(d,p)	1.55 (0.05)	1.62 (0.12)	1.57 (0.07)	**1.51 (0.01)**	**1.47 (-0.03)**	**1.50 (0.00)**	**1.51 (0.01)**
TZP	1.57 (0.07)	1.65 (0.15)	1.61 (0.11)	1.55 (0.05)	**1.52 (0.02)**	**1.53 (0.03)**	1.55 (0.05)
cc-pVTZ	1.54 (0.04)	n/a	1.57 (0.07)	**1.51 (0.01)**	**1.48 (-0.02)**	**1.50 (0.00)**	**1.52 (0.02)**

Table 7. Calculated dipole moments for oxazole (**21**) in Debye units with deviation from experimental value (1.50 ± 0.03) shown in parentheses. Values within experimental deviation shown in bold type.

In addition to the clear trend observed for the level of theory used, those basis sets which provided the most results within the experimental range were easily seen. These were the 6-31G – based double-ζ sets and the triple-ζ 6-311G(d,p) and cc-pVTZ sets. These sets uniformly provided values that were within experimental uncertainty with all of the gradient-corrected density functionals. With the exception of MP2, all other values calculated with these basis sets were reasonably close to the experimental value of 1.50 Debye. The poor value given by the BP86/6-311G(d) calculation (deviation = -0.33 Debye) was thought to have been a mistake but when the original output file was checked this was found to be correct.

Basis	RHF	MP2	SVWN	BP86	BLYP	HCTH	B3LYP
Double-ζ							
6-31G(d)	1.45 (-0.16)	1.72 (0.11)	1.44 (-0.17)	1.38 (-0.23)	1.33 (-0.28)	1.39 (-0.22)	1.43 (-0.18)
6-31G(d,p)	1.44 (-0.17)	1.73 (0.12)	1.44 (-0.17)	1.38 (-0.23)	1.33 (-0.28)	1.39 (-0.22)	1.43 (-0.18)
DZP	1.53 (-0.08)	1.87 (0.26)	**1.63 (0.02)**	1.54 (-0.07)	1.57 (-0.04)	1.54 (-0.07)	**1.62 (0.01)**
cc-pVDZ	1.33 (-0.28)	**1.61 (0.00)**	1.40 (-0.21)	1.33 (-0.28)	1.31 (-0.30)	1.34 (-0.27)	1.39 (-0.22)
SBKJC(d)	1.53 (-0.08)	1.86 (0.25)	**1.60 (-0.01)**	1.54 (-0.07)	1.49 (-0.12)	1.56 (-0.05)	1.56 (-0.05)
SBKJC(d,p)	1.53 (-0.08)	1.88 (0.27)	**1.60 (-0.01)**	1.54 (-0.07)	1.49 (-0.12)	1.56 (-0.05)	1.56 (-0.05)
Triple-ζ							
6-311G(d)	1.39 (-0.22)	1.69 (0.08)	1.48 (-0.13)	1.40 (-0.21)	1.33 (-0.28)	1.34 (-0.27)	1.44 (-0.17)
6-311G(d,p)	1.39 (-0.22)	1.70 (0.09)	1.49 (-0.12)	1.40 (-0.21)	1.33 (-0.28)	1.35 (-0.26)	1.44 (-0.17)
TZP	1.48 (-0.13)	1.85 (0.24)	**1.63 (0.02)**	1.55 (-0.06)	1.55 (-0.06)	1.53 (-0.08)	**1.60 (-0.01)**
cc-pVTZ	1.42 (-0.19)	n/a	1.55 (-0.06)	1.47 (-0.14)	1.46 (-0.15)	1.44 (-0.17)	1.52 (-0.09)

Table 8. Calculated dipole moments for thiazole (**22**) in Debye units. Deviation from experimental value (1.61 ± 0.03) shown in parentheses. Values within experimental deviation shown in bold type.

Contrary to the trends observed in the calculated oxazole (**21**) dipole moments, the corresponding calculations on thiazole (**22**) showed that the best overall performance for this molecule was given by the SVWN functional with four of its values falling within the experimental range (1.61 ± 0.03 Debye). B3LYP gave two values within this range and MP2 gave one (the MP2/cc-pVDZ value was the only one obtained that had a zero deviation). None of the GGA functionals or the RHF method gave values with deviations ≤ 0.03 Debye.

Some of the lowest deviations outside the experimental range were found when the SBKJC basis sets were combined with the DFT functionals or the RHF method. These basis sets did not produce particularly good results when combined with the MP2 approach. With the exception of the BP86/TZP, BLYP/TZP and B3LYP/cc-pVTZ combinations, the triple-ζ quality basis sets gave no significant improvement over the double-ζ sets and their use frequently lead to increased deviations from experiment.

Harmonic vibrational frequencies

The results of the vibrational frequency calculations using the 6-31G(d,p) basis are given in Tables 9 (oxazole (**21**)) and 10 (thiazole (**22**)). The mean scaling factor (MSF) required to bring the values for each method into line with those determined experimentally are given at the bottom of each column.

For the vibrational frequencies of oxazole (**21**) the MSF values showing the smallest deviation from unity were those obtained with the SVWN (MSF=0.993), BP86 (MSF=1.009) and HCTH (MSF=0.991) functionals. BLYP (MSF=1.015) and B3LYP (MSF=0.973) were next in the sequence with MP2 (MSF=0.964) and RHF (MSF=0.886) requiring the largest degree of scaling in order to bring their calculated frequencies into line with experiment. Thus, the order of decreasing accuracy in these calculations was:

$$\text{SVWN} > \text{HCTH} = \text{BP86} > \text{BLYP} > \text{B3LYP} > \text{MP2} >> \text{RHF}$$

The corresponding sequence for thiazole (**22**) was:

$$\text{HCTH} \approx \text{BP86} > \text{BLYP} > \text{SVWN} \approx \text{B3LYP} > \text{MP2} >> \text{RHF}$$

The main (qualitative) difference between the results for the two molecules was the change in the order of accuracy that placed the SVWN functional below the GGA functionals in the sequence. Quantitatively, all the methods except SVWN required scaling factors for the thiazole frequencies that were almost identical (in their deviation from unity) to those required for oxazole (see Tables 9 and 10). The deviation from unity in the MSF required for the SVWN thiazole frequencies (+0.023) was approximately three times larger and of opposite sign to that found in the corresponding oxazole frequencies (-0.007).

Vibration		Exp.	RHF	MP2	SVWN	BP86	BLYP	HCTH	B3LYP
A'	1	3170	3486.3	3393.8	3237.3	3231.2	3232.1	3286.6	3326.4
	2	3144	3453.5	3371.9	3213.7	3203.8	3198.9	3256.1	3294.2
	3	3141	3445.6	3356.2	3196.3	3191.0	3191.4	3239.5	3279.6
	4	1537	1778.0	1585.7	1562.0	1528.9	1520.4	1558.3	1593.3
	5	1504	1713.5	1548.8	1520.7	1488.6	1477.0	1518.5	1544.6
	6	1324	1506.0	1392.9	1354.8	1321.1	1312.5	1349.8	1370.5
	7	1252	1399.7	1303.6	1223.0	1235.3	1241.9	1256.5	1281.5
	8	1139	1298.0	1213.3	1181.9	1130.8	1124.7	1155.1	1174.9
	9	1086	1226.6	1159.6	1164.6	1104.5	1082.2	1132.1	1132.7
	10	1078	1192.8	1136.4	1083.9	1062.9	1041.9	1086.7	1107.6
	11	1046	1162.5	1105.5	1078.7	1048.3	1032.8	1072.6	1082.7
	12	899	1014.4	913.5	893.0	882.0	881.6	893.8	914.9
	13	854	996.5	845.5	830.6	824.0	822.4	838.6	867.7
A"	14	907	1020.9	921.6	907.9	891.7	887.7	906.5	925.8
	15	830	993.2	818.2	779.7	779.4	778.5	795.1	828.0
	16	750	878.0	747.7	721.6	716.2	713.2	726.0	750.3
	17	647	711.4	668.7	659.8	647.4	644.7	657.2	664.7
	18	607	671.8	625.1	624.1	608.0	603.6	617.6	625.4
Mean Scale Factor			0.886	0.964	0.993	1.009	1.015	0.991	0.973

Table 9. Harmonic vibrational frequencies for oxazole (**21**) calculated with the 6-31G(d,p) basis set. Experimental values from El-Azhary, 1995.[33]

Vibration		Exp.	RHF	MP2	SVWN	BP86	BLYP	HCTH	B3LYP
A'	1	3140	3440.9	3352.9	3239.1	3197.8	3193.6	3246.4	3277.4
	2	3093	3411.6	3327.4	3198.8	3158.2	3153.5	3212.8	3242.2
	3	3065	3405.1	3311.8	3196.7	3155.4	3150.7	3200.0	3235.1
	4	1484	1735.8	1522.8	1470.9	1483.6	1477.4	1509.7	1545.3
	5	1383	1617.6	1439.3	1402.9	1413.2	1399.1	1438.9	1458.3
	6	1325	1490.0	1394.6	1281.8	1320.4	1316.2	1345.4	1365.7
	7	1239	1392.5	1298.5	1185.2	1224.8	1231.3	1246.4	1272.1
	8	1124	1245.4	1183.1	1102.7	1129.6	1125.5	1149.8	1161.4
	9	1043	1155.4	1096.5	1007.7	1040.3	1035.7	1060.4	1074.2
	10	888	1047.8	921.4	840.2	857.0	852.8	874.5	902.3
	11	867	969.0	907.1	831.6	849.6	851.1	866.2	886.1
	12	759	834.6	788.8	697.3	723.6	702.2	741.6	744.8
	13	612	669.2	635.3	613.6	595.0	592.0	605.0	615.3
A"	14	849	946.1	878.5	819.1	837.3	824.3	854.5	864.1
	15	798	926.3	801.2	733.9	764.8	764.2	777.4	804.1
	16	717	814.0	727.0	669.8	694.2	696.4	706.6	729.9
	17	603	659.8	612.9	576.7	593.0	582.1	604.1	612.6
	18	467	502.5	474.9	480.3	459.6	455.4	467.4	471.3
Mean Scale Factor			0.892	0.961	1.023	1.010	1.016	0.992	0.976

Table 10. Harmonic vibrational frequencies for thiazole (**22**) calculated with the 6-31G(d,p) basis set. Experimental values from El-Azhary, 1995.[33]

Summary of DFT results for oxazole (21) and thiazole (22)

Overall, the DFT approach to the calculation of the equilibrium ground state geometries of oxazole (**21**) and thiazole (**22**) showed good agreement with the gas-phase MW geometries obtained from the literature.[37] The best performance was given by the HCTH (GGA) and B3LYP (HF-DFT) functionals which were able to reproduce the bond lengths of both molecules at least as well as the MP2 wavefunction-based method and frequently gave superior results. Both DFT methods had the added advantage of requiring considerably less CPU time than MP2. Despite being the simplest of the functionals tested, SVWN also performed well with mean absolute deviations for bond lengths of 0.01 Å or less when all-electron basis sets were used. The GGA functional BP86 performed similarly to the uncorrelated RHF method but showed some improvement over this method when the largest basis sets were used. The worst overall performance was given by the BLYP (GGA) functional with deviations that were often two or three times larger than those of HCTH or B3LYP.

For bond angles the trend was less clear, with all of the methods tested giving reasonably similar MAD values. The worst performance was again that of the BLYP functional, which gave deviations that were approximately 1.5 – 2.0 times larger than those of the other DFT methods. Overall, SVWN outperformed the remaining functionals (BP86, HCTH and B3LYP) which all gave results comparable to those obtained at the MP2 level.

With the all-electron basis sets, the effect on the accuracy of the results obtained from the geometry optimizations of varying the size and/or valence space quality of the basis was negligible. Significantly, the standard double-ζ 6-31G basis was seen to give some of the most accurate results for all of the methods tested here. The use of triple-ζ basis sets generally gave no improvement over the smaller double-ζ sets. Relative to the all-electron sets, the SBKJC ECP basis sets gave noticeably worse bond lengths overall. However, the bond angles obtained with SBKJC were extremely competitive for both molecules.

In the calculation of the dipole moments of **21** and **22** the HCTH, B3LYP and BP86 functionals gave equally good performances. For oxazole, BLYP actually provided the

largest number of values within the experimental range but failed to provide any results of this accuracy in the calculations on thiazole. The average performance of the DFT methods for this molecular property was similar to, or slightly better than that of the RHF calculations and better than MP2. Basis set effects in these calculations were again small with the 6-31G sets proving to be extremely competitive (particularly in the case of oxazole where these sets gave results within experimental accuracy with all of the functionals containing gradient corrections).

All of the mean scaling factors (MSF) required by the DFT methods to bring the calculated vibrational frequencies of **21** and **22** into line with experiment showed lower deviations from unity than those of either MP2 or RHF. The magnitude of the MSF required by each DFT method was almost identical for the two molecules, with the exception of SVWN which required a much larger MSF for thiazole than for oxazole.

Discussion

The results of this study into the performance of various DFT methods show that, in general, the best reproduction of the experimental geometries of oxazole (**21**) and thiazole (**22**) is obtained when corrections for the (local) gradient of the charge density are included in the approximate exchange-correlation functional. An exception to this rule is found in the results obtained with the BLYP GGA functional which gave the poorest performance of all the DFT methods and was similar to the uncorrelated RHF method. The local density approximation (which neglects gradient effects) also performs reasonably well, particularly in the calculations on thiazole. Of the functionals tested, HCTH and B3LYP perform at least as well as the MP2 post-HF method.

The fact that the bond lengths in thiazole (**22**) show significantly larger deviations from experiment than those of oxazole (**21**) can be related to a less accurate description of the C-S bonds in **22** than the C-O bonds in **21**. The presence of 2nd row main group atoms is known to reduce the accuracy of theoretical calculations of this type compared to calculations on the equivalent 1st row species.[39] The fact that none of the methods stood out particularly in the calculation of bond angles is probably a reflection on the constraints placed on these angles by the cyclic nature of the compounds studied.

Within the range of basis sets used in the present geometry optimizations, the DFT methods do not generally appear to require the largest sets in order to provide their best results. In fact, the results of this study seem to suggest that going beyond the 6-31G(d,p) basis is likely to lead to a deterioration on accuracy when calculations are performed on molecules of this type. A small improvement is seen when the largest of all the sets tested here (cc-pVTZ) is used, in line with the observations of El-Azhary and Suter, 1996.[37] The double-ζ equivalent of this basis (cc-pVDZ) requires considerably more CPU time than the 6-31G sets but the results obtained with this basis set are of relatively poor quality and do not justify the additional computational expense (this is also in agreement with the findings of El-Azhary and Suter). Of the triple-ζ sets, those based on 6-311G did not produce the improvement over the 6-31G sets that might be expected. The reason for this may be related to the observation by Grev and Schaefer, 1989,[50] that the orbital (ζ) exponents used in the formulation of this basis are too tight with the result that most of the functions intended to describe the valence region tend to 'fall' into the core, resulting in a basis that is more properly labelled '63-11G'.

The results obtained with the relativistic effective core potential SBKJC basis set included in this study are clearly in significantly worse agreement with the MW values for the geometry of the two compounds investigated than any of those found with the all-electron sets. However, it is important to bear in mind that this is a relative observation and that objectively, the ECP results are actually quite reasonable considering the saving in CPU time that can be achieved when using basis sets of this type. This saving is not, perhaps, particularly great in calculations on small main group compounds but if larger systems, or systems containing heavier atoms such as the transition metals are the target of a study, then the reduction in accuracy with these sets may be offset by the ability to actually complete the calculations. In addition, the level of accuracy that can potentially be obtained with some of the large all-electron basis sets may significantly exceed that of the available experimental data for compounds of this type.

An apparent basis set insensitivity of dipole moments calculated using DFT has been commented on previously.[40] This is supported in general by the performance of the DFT

methods tested here and in particular by the results of the BLYP calculations on oxazole (**21**). Of the remaining DFT methods, the BP86, HCTH and B3LYP functionals give a similar degree of accuracy for this property in both **21** and **22**. SVWN does not behave reliably with the molecules studied here as it gives the best overall performance of the DFT methods for **22** but gives the worst performance for **21**. Considering the relatively low level of sophistication of this functional, the accuracy of the thiazole results may have been fortuitous and probably do not reflect the quality of the physical description of the molecule provided by this method.

With respect to the magnitude of the mean scale factors required for the harmonic vibrational frequencies obtained in this work, the comments made above concerning the erratic behaviour of the SVWN functional in the geometry optimizations and dipole moment calculations seem to also be applicable to the calculation of this property. Other authors have observed this system dependence when using the SVWN method, suggesting that this might be a general feature of the local density approximation.[39,51] A systematic study of a number of 1st row main group molecules with six density functionals in combination with the 6-31G(d) basis by Zhou and co-workers, 1996, highlighted a general overestimation of fundamental frequencies with the B3LYP, B3P86 and BHLYP hybrid functionals.[39] SVWN, BLYP and BP86 gave slightly larger average deviations. The fact that the trend observed in the present work does not agree with this last report is most likely a consequence of the cyclic nature of compounds **21** and **22** as none of the molecules studied by Zhou *et al*. (*e.g.*, ethylene, formaldehyde and butadiene) contained this feature in their structures and can therefore be expected to display different dynamical properties.

The good overall performance of the B3LYP functional[41] is not, perhaps, very surprising as this is one of the most commonly used of the more recent functional forms and the quality of the results obtained with this method are accepted as being similar to the highly sophisticated coupled-cluster singles, doubles and non-iterative triples (CCSD(T)) post-HF method and generally as good as, or better than, the MP2 method.[39,40] This has been attributed to the inclusion of a fraction of 'exact' Hartree-Fock exchange, as the BLYP functional (which differs from B3LYP in that it lacks the HF exchange contribution) gives

much lower accuracy.[39] The combination of this type of hybrid exchange with other correlation functionals has also been shown previously to provide an improvement over the corresponding 'pure' DFT approach and for this reason a general consensus has arisen in the literature suggesting that these hybrid models represented the way forward in attempts to improve the performance of density functional methods. The excellent all-round performance of the B3LYP functional has lead to it being forwarded as an alternative to the HF method in the calculation of geometries within the highly accurate G2 and CBS computational schemes.[52,53]

The generalized gradient approximation HCTH functional of Hamprecht *et al*., 1998,[46] is representative of the latest generation of 'pure' exchange-correlation functionals and contains no admixture of Hartree-Fock exchange. With this in mind, the excellent performance of this functional in replicating the experimental geometries of the two test molecules is particularly interesting as it suggest that no 'dilution' with components of a separate theory is required for DFT to perform as well as the much more time-consuming post-HF wavefunction based methods. This was the amongst the original conclusions of the creators of this functional when they compared HCTH with MP2 and several hybrid HF-DFT functionals.[46] The training set used by these authors in the development of the functional contained no cyclic molecules similar to either of the systems studied in the present work but in a subsequent paper Altmann and Handy, 1999,[47] studied a number of sulfur-containing systems with the HCTH functional and compared the results with those obtained with the MP2 and BLYP methods. The authors again found that HCTH outperformed both of these established methods.

In conclusion, the combinations of the HCTH and B3LYP functionals with the polarized 6-31G basis set provide results for the geometries of oxazole (**21**) and thiazole (**22**) that are in excellent agreement with experiment and are of comparable quality to the best MP2 results and to equivalent calculations with the other density functionals tested. This is an important finding with respect to the intended use of DFT in the studies on the full patellamide molecules, as the four subunits of these peptides based on the heterocycles **21** and **22** constitute a considerable fraction of the total structures and also appear to be of major importance in determining their conformational properties. The size of the

patellamides (~100 atoms in total, approximately half of which come from the main group of the periodic table) means that keeping the number of basis functions to a minimum is of utmost importance. The present study suggests that this might be achieved by using a polarized 6-31G basis at no significant cost in terms of accuracy.

Chapter 2

Density Functional Investigation of Equilibrium Geometries and Conformational Energetics of Ascidiacyclamide and Patellamides A, C and D

Introduction

After attempting several trial calculations on the full patellamide structures it was realised that performing a comparative theoretical study into the conformational preferences displayed by the patellamides within a reasonable amount of time was beyond the capabilities of the systems available at Aberdeen. Fortunately Dr. Martyn Guest at the CLRC's Daresbury Laboratory kindly agreed to assist in these studies by running calculations on the institute's IBM-SP supercomputer and linux-based Alpha cluster using the parallel implementation of GAMESS-UK. As most of this work took place over the summer period, these systems were relatively free of users and consequently it was possible to perform a large number of geometry optimizations on several members of the patellamide family.

Methods

The level of theory selected for use in these studies involved the use of the B3LYP density functional in combination with the 6-31G(d,p) basis set. This combination had been seen to provide a high degree of accuracy previously in the investigations into the structurally relevant heterocycles oxazole (**21**) and thiazole (**22**). Furthermore, several workers had used this level of theory in investigations into the geometries and energetics of polypeptide systems and found good agreement with experiment.[54-58]

The HCTH functional had also provided results of high quality in the study on **21** and **22** but as no examples of its use in large systems could be found in the literature its applicability could not be evaluated and consequently it was not selected for use in this study. The patellamides selected for inclusion in this study (Figure 9) were the C_2-symmetric ascidiacyclamide (ASC; **1**) and the variously substituted patellamides A (PatA; **2**), C (PatC; **4**) and D (PatD; **5**). These were felt to be representative of the differing conformational preferences shown by the patellamide family. Overall, the preference shown by ASC and PatA is for the open conformation whilst PatC and PatD are found in the closed conformation. Compound **1** displays C_2 symmetry with **2** only varying from this by the lack of a single methyl group on one of the oxazoline units. In **5** the two D-Val residues have been replaced by one D-Phe and one D-Ala. This is also seen in **4** but in this case one of the L-Ile residues has also been replaced by a L-Val. Thus, the pairs **1** & **2**

and **4** & **5** differ considerably from one another but there is only a difference of one CH_3 unit between the members of a pair.

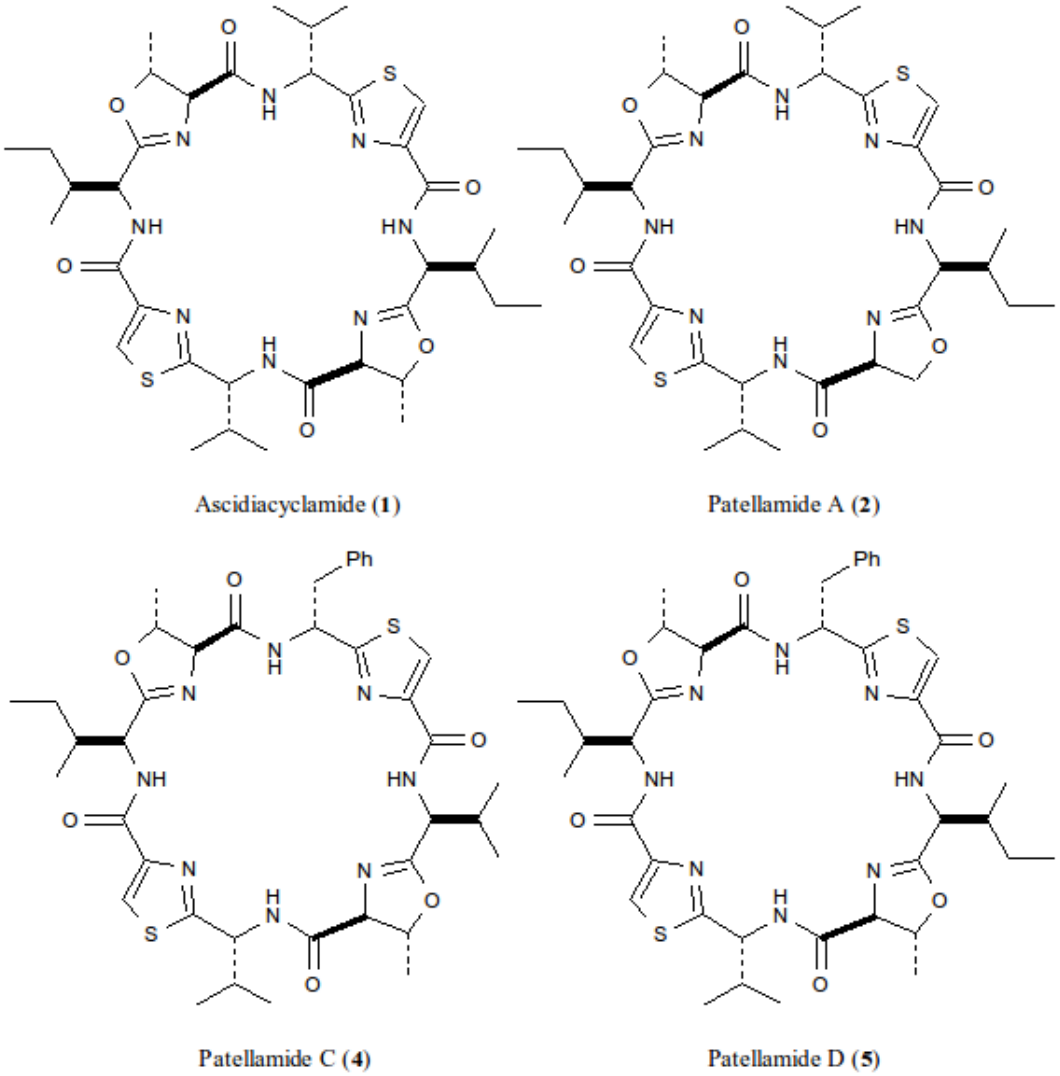


Figure 9. Structures of members of the patellamide family chosen for inclusion in the DFT study.

The reason for including two from each overall conformational family was that there was some evidence from CD studies that PatA could be induced to adopt the closed conformation through reductions in temperature[19] or changes in solvent polarity[17] and so it was decided to compare the conformational energetics of this molecule with those of ASC as these systems differed by only a single methyl group on one of the oxazoline rings and the latter molecule had not been observed in the 'closed' state. PatC and PatD were likewise selected as they also only differed by one methylene unit and it was of interest to

find out if an apparently insignificant alteration such as this could lead to noticeable effects on the structures and energetics of the molecules as a whole.

Initial geometries for ASC and PatA in the open conformation were obtained from the published XRD structures.[14,26] These were modified using the model-building feature in the Macromodel molecular mechanics package (v6.5)[59] to produce the corresponding structures for PatC and PatD. Similarly, the XRD structure of PatD was used to produce the closed conformations of ASC, PatA and PatC. These structures, and several of their side-chain rotamers, were then optimized with the MM2* force field prior to their use in the B3LYP calculations. Examples of the rotamer types for which the DFT calculations were performed are shown in Figures 10 & 11. The energy differences observed in these calculations were used as a rough guide in the selection of structures that were likely to be significantly populated at room temperature. Based on their MM2* energies, four rotamers each were selected in total for ASC and PatA. Because of similarities in the energies of several of the phenylalanine rotamers of PatC and PatD a total of five structures were selected for each of these compounds.

In the DFT calculations, an initial optimization using the unpolarized 6-31G basis set was performed in order to reduce the overall time required. This was followed by refinement of these intermediate geometries with the full 6-31G(d,p) basis. The number of functions required for each molecule are shown in Table 11. This two-step procedure required approximately 600 functions in the initial stages whereas during the refinements around 1000 functions were required to describe each of the patellamides.

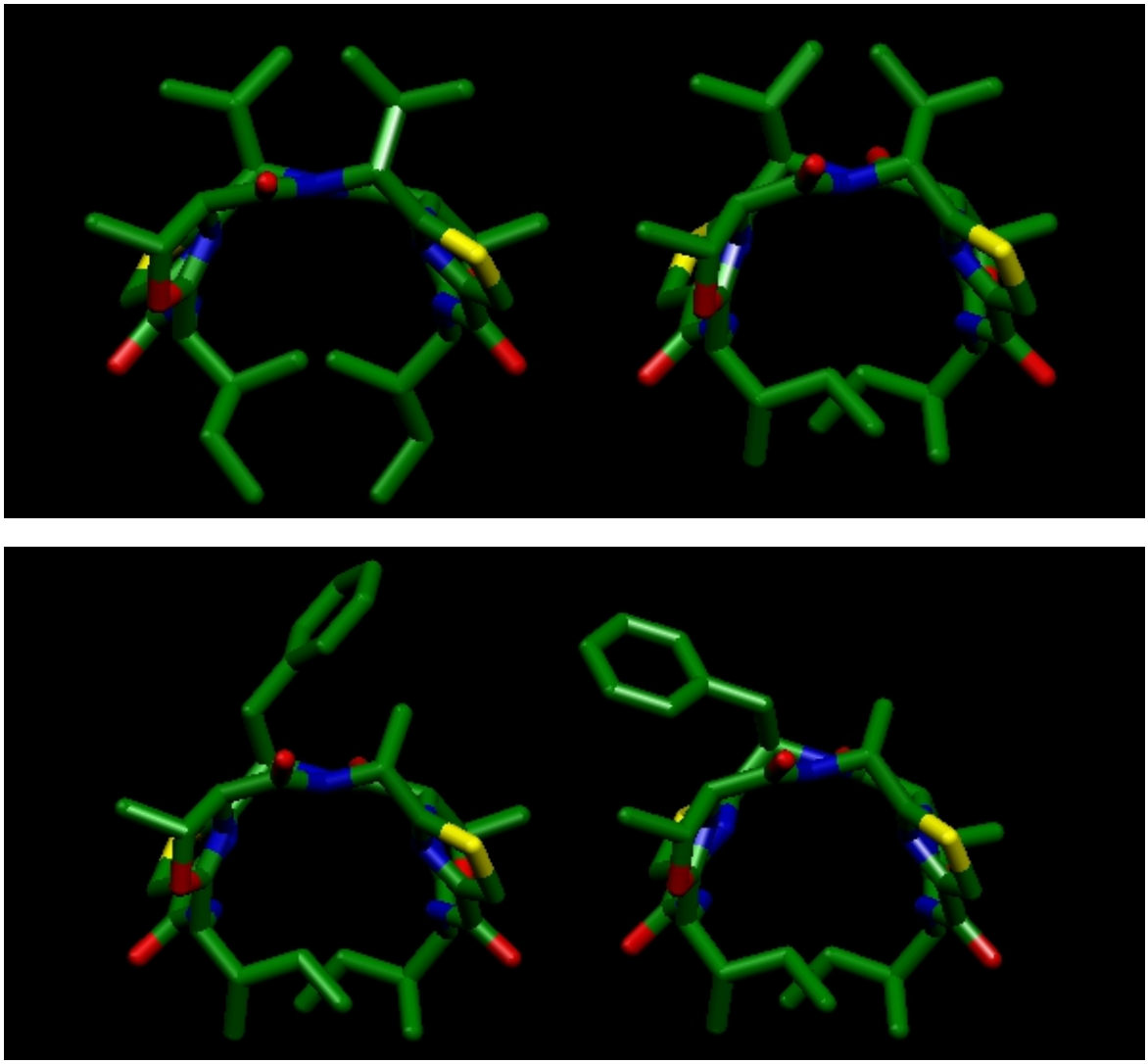


Figure 10. *Top:* End view of ascidiacyclamide open (overall) conformation showing isoleucine rotamers selected for optimization at DFT level. These rotamers were also selected for patellamides A, C and D. *Bottom:* End view of patellamide D open (overall) conformation showing phenylalanine rotamers selected for optimization at DFT level. These rotamers were also selected for patellamide C. The member of each pair having lowest energy is shown on the right of the page. Hydrogens omitted to improve clarity.

As 18 of these structures were optimized in total this represented a significant reduction in the net time required for these calculations, especially as the initial stages were expected to be more time-consuming than the final points in the optimizations. Because of the way that the computer systems were administered at Daresbury, individual jobs were limited to four hours and consequently the calculations on the patellamides had to be restarted several times. In total, each of the 18 structures required approximately 48 hours of CPU time when running on all processors of either the IBM-SP or the Alpha-cluster.

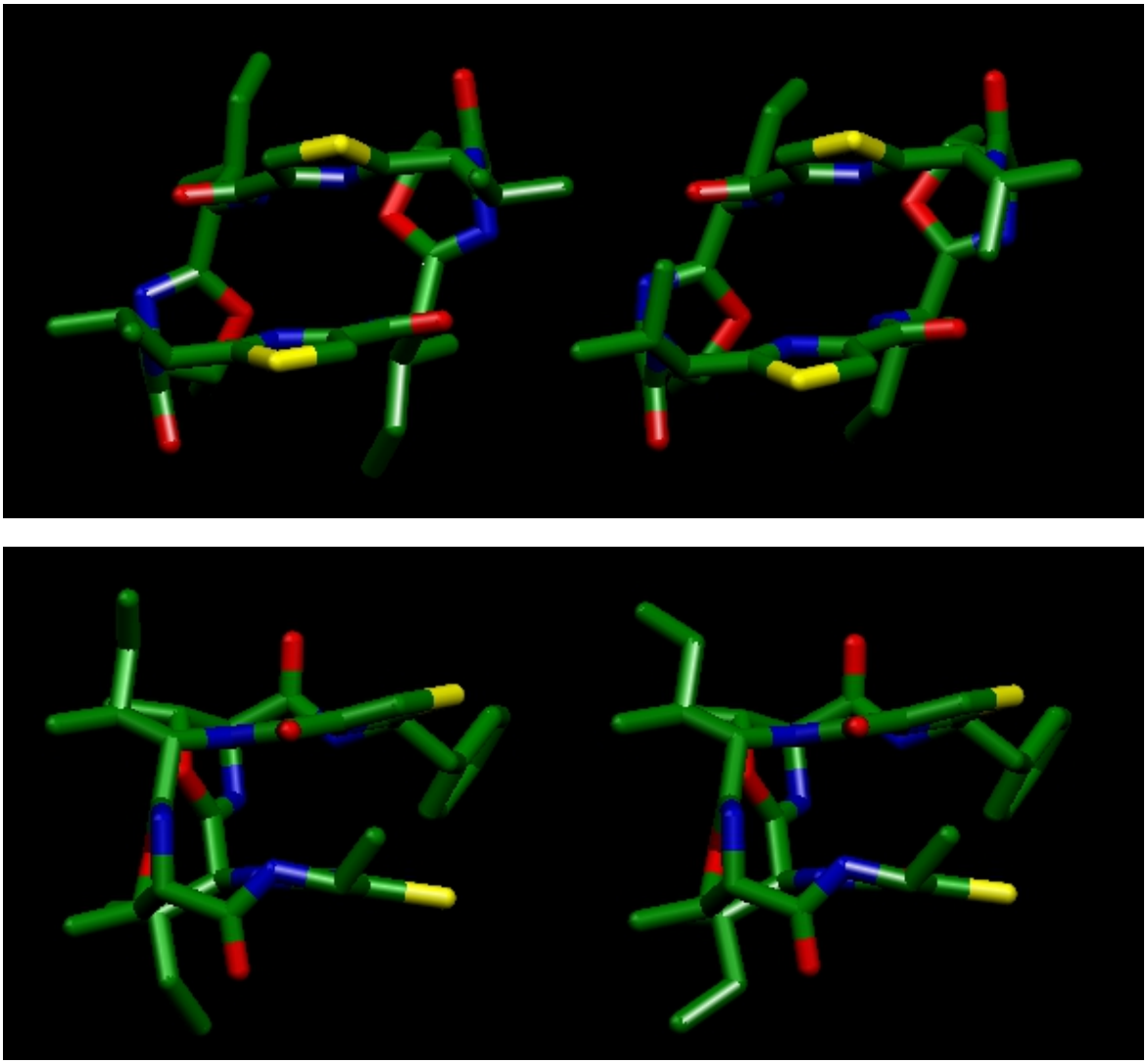


Figure 11. *Top:* Front view of ascidiacyclamide closed (overall) conformation showing valine rotamers selected for optimization at DFT level. These rotamers were also selected for patellamide A. *Bottom:* End view of patellamide D closed (overall) conformation showing isoleucine rotamers selected for optimization at DFT level. These rotamers were also selected for patellamide C. The member of each pair having lowest energy is shown on the right of the page. Hydrogens omitted to improve clarity.

No attempt to make use of symmetry during these calculations was made as this would possibly have interfered with the course of the optimizations. The rationale behind this decision was that although the previous experimental work on these compounds had focussed to a large extent on the substitutional symmetry of the molecules, the effects of substitution on the backbone symmetry had generally been neglected (some data on this feature could be extracted from the XRD structures but, as mentioned above, the variability displayed by this method meant that any conclusions would be difficult to qualify). Consequently it was hoped that as a result of the neglect of such environmental

effects in the present work it might be possible to quantify any changes of this type by comparing the C_2-symmetric ASC with the other compounds in the study. The initial ASC input structure was deliberately desymmetrized by rotating the amino acid side chains by varying (small) amounts yielding an overall C_1 symmetry. Initial optimization with the unpolarized basis set produced an asymmetric peptide backbone but the C_2 symmetry was regained after refinement of the structure at the B3LYP/6-31G(d,p) level.

Peptide	6-31G	6-31G(d,p)
Ascidiacyclamide (**1**)	580	1048
Patellamide A (**2**)	567	1023
Patellamide C (**4**)	577	1033
Patellamide D (**5**)	590	1058

Table 11. Number of basis functions required by each of the patellamides in the DFT study.

Results

Conformational energetics

The lowest energy side-chain rotamers were selected for each of the peptides in both the open and closed conformations and their energies compared. For all of the peptide conformers studied the second lowest energy side chain rotamers lay approximately 10 KJ Mol^{-1} above the lowest energy structures suggesting that their populations would be negligible at room temperature (assuming thermodynamic equilibrium) and as a result, these and the other higher energy structures were not considered further. The lowest conformational energy differences are shown in Table 12 along with the energy differences obtained for the corresponding structures using molecular mechanics. The force fields used in these calculations were MM2* and AMBER*, both of which are provided with the Macromodel package.

The molecular mechanics (MM) methods both showed large negative energy changes on going from the open to the closed conformations for all four peptides. B3LYP/6-31G gave negative changes for PatC and PatD whilst the changes for ASC and PatA were positive. This result appears to agree with the trend observed in the experimental studies as from these values ASC and PatA would be expected to prefer the open conformation with PatC and PatD preferring the closed conformation. On addition of polarization functions all of

the energy changes became positive, however, as can be seen in Table 12 the ΔE values for PatC and PatD remained considerably lower than those of either ASC or PatA.

Peptide (preferred conformation)	B3LYP/6-31G	B3LYP/6-31G(d,p)	MM2*	AMBER*
ASC (open)	14.90	21.40	-59.81	-57.00
PatA (open)	10.60	16.40	-60.74	-59.98
PatC (closed)	-4.00	5.80	-67.47	-52.09
PatD (closed)	-6.00	3.60	-72.28	-47.33

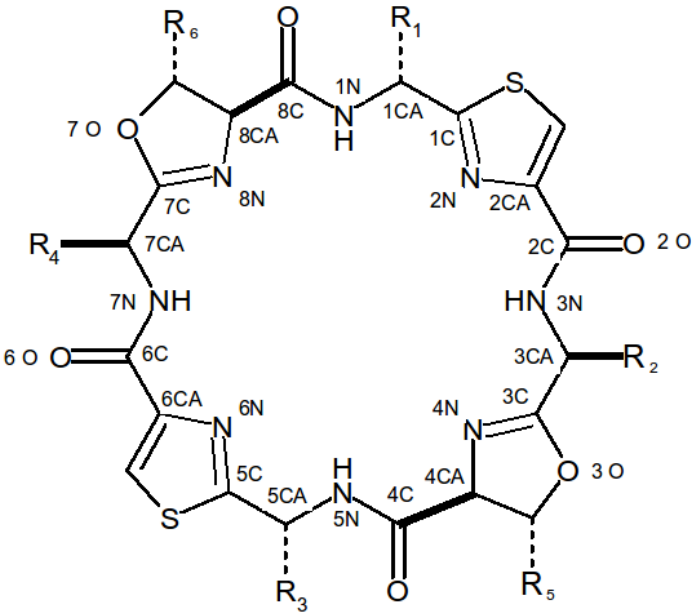
Table 12. Conformational energy differences (open→closed) for ASC (**1**), PatA (**2**), PatC (**4**) and PatD (**5**). All values in KJ mol^{-1}.

The MM energies given by either of the force fields do not (quantitatively or qualitatively) reproduce the well-defined trend observed in the B3LYP calculations. The DFT results show a reduction in ΔE of approximately 30% (6-31G) and 25% (6-31G(d,p)) on going from ASC to PatA and approximately 60% (6-31G) and 75% (6-31G(d,p)) between ASC and the phenylalanine-containing peptides PatC/D. The MM2* method correctly predicts the qualitative difference between the ASC/PatA and PatC/PatD pairs with the latter having the lowest conformational energies but fails to place the members of each pair in the order given by B3LYP/6-31G or B3LYP/6-31G(d,p), with ASC and PatC having the lowest ΔE values out of each pair. AMBER* provides even worse agreement with the B3LYP results, reversing the order of the pairs and of the members of each pair. The fact that the DFT results obtained in this study are not susceptible to the uncertainties inherent in the MM methods (*e.g.*, the limited transferability of force field parameters between different molecular environments or the application of parameters derived from crystal structures to gas phase calculations) suggests that neither the MM2* or AMBER* force fields can reliably be used in the gas phase prediction of the conformational preferences of the patellamides.

In the B3LYP results, the large difference between PatC/D and ASC was not unexpected in light of the large difference in the substitution patterns of these molecules. The ~25% reduction in ΔE on going to PatA, however, can only be attributed to the effects of loss of the methyl group from one of the oxazoline rings and supports the previously suggested importance of these structural subunits in determining the conformational properties of the

patellamides. The fact that a similar (if slightly smaller) reduction occurs between PatD and PatC was somewhat surprising as the difference between these peptides is merely the loss of a CH_2 unit adjacent to the terminal methyl group in one of the L-amino acid side chains.

Peptide backbone (macrocycle) geometries



(**1**) Ascidiacyclamide R_1=R_3=D-Val R_2=R_4=L-Ile R_5=R_6=CH_3
(**2**) Patellamide A R_1=R_3=D-Val R_2=R_4=L-Ile R_5=H R_6=CH3
(**4**) Patellamide C R_1=D-Phe R_2=L-Val R_3=D-Ala R_4=L-Ile R_5=R_6=CH_3
(**5**) Patellamide D R_1=D-Phe R_2=R_4=L-Ile R_3=D-Ala R_5=R_6=CH_3

Figure 12. Numbering scheme used in description of the patellamide backbone structure.

The B3LYP/6-31G(d,p) macrocycle geometries of the four peptides in both conformations were studied and both forms of ASC were found to have almost perfect C_2 symmetry with no deviations from this seen in measurements accurate to 5 decimal places for bond lengths (Å) and 3 decimal places in bond angles and dihedrals (º). Because of this excellent agreement between the two halves of ASC combined with the fact that no symmetry constraints had been applied it was decided that comparisons of the different patellamides could confidently be made to 3 decimal places in the bond lengths and 1 decimal place in bond angles (a commonly quoted accuracy in the previous experimental studies). Dihedral angles were expected to be relatively flexible and so measurement and comparison of this coordinate type was arbitrarily restricted to an accuracy of 1º. The

lengths, angles and dihedrals found for ASC were used for comparison with the corresponding geometric coordinates in the other three patellamides in order to assess the deviation from C_2 symmetry caused by the various substitutions relative to ASC present in these molecules. The atom numbering system used in these comparisons is shown in Figure 12. The deviations observed for bond lengths, angles and dihedrals are shown graphically in Figures 13-18 with spline functions fitted to the data in order to highlight the overall trends in the deviations.

The deviations in the bond lengths (Figure 13) were all small in magnitude and consequently it was not felt that they could provide a great deal of useful information individually. However, by viewing the peptide backbone as a whole it was possible to observe some trends relating to substitutions in amino acid residues. For PatA in the open conformation deviations from the ASC geometry were seen to be almost exclusively clustered in residues 3-5 containing the demethylated oxazoline ring.

Inspection of the graph for bond lengths in the closed conformation (Figure 14) reveals that the folding of PatA caused a reduction in the degree of asymmetry, with the number of deviations observed being reduced by half and spread much more equally throughout the structure. The largest deviation (0.002 Å) observed in PatA occurred in residue 4, regardless of conformation.

PatC and PatD displayed almost identical patterns of bond length deviations in the open conformation although a small number of differences were evident. These were mainly of the order of 0.001 Å but at residue 3 PatC displayed a deviation from ASC that was twice as large (0.006 Å) as that of PatD. This large deviation is perhaps to be expected for PatC (residue 3 = L-Val as opposed to L-Ile in ASC), as might the deviations around the D-Phe and D-Ala residues (1 and 5) of both PatC and PatD.

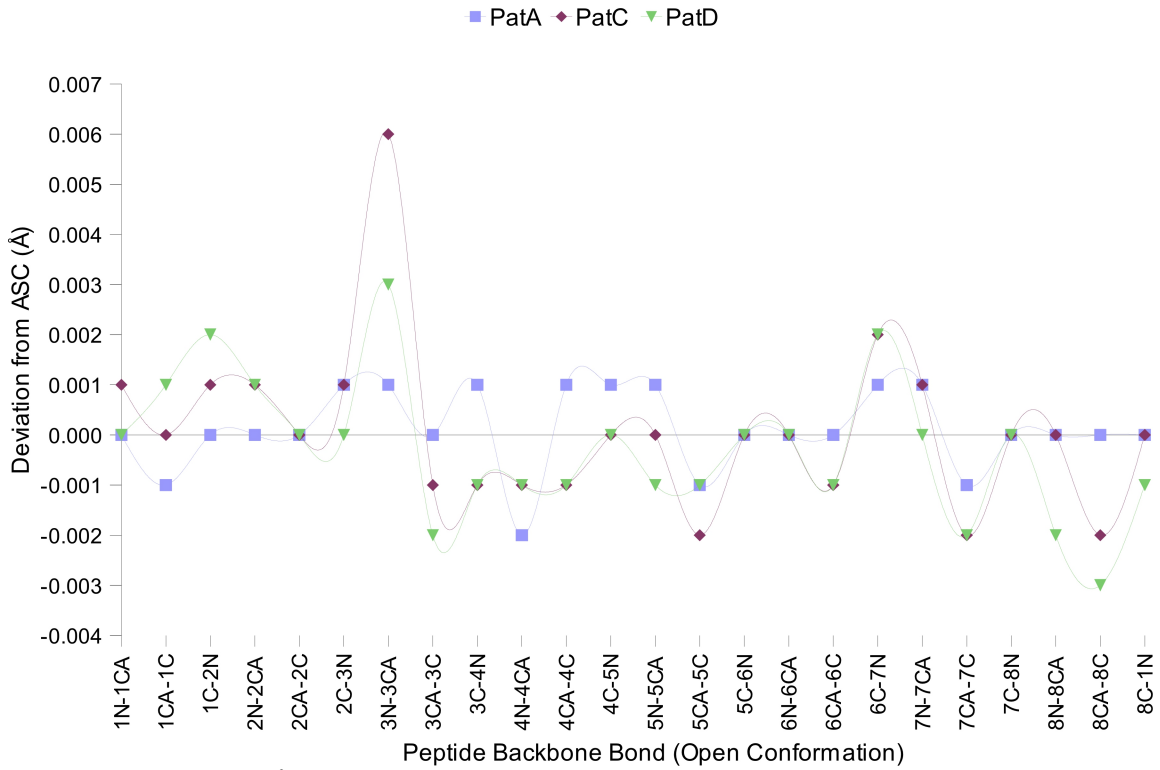


Figure 13. Deviations in Å from ASC bond lengths (open conformation). B3LYP/6-31G(d,p) geometries.

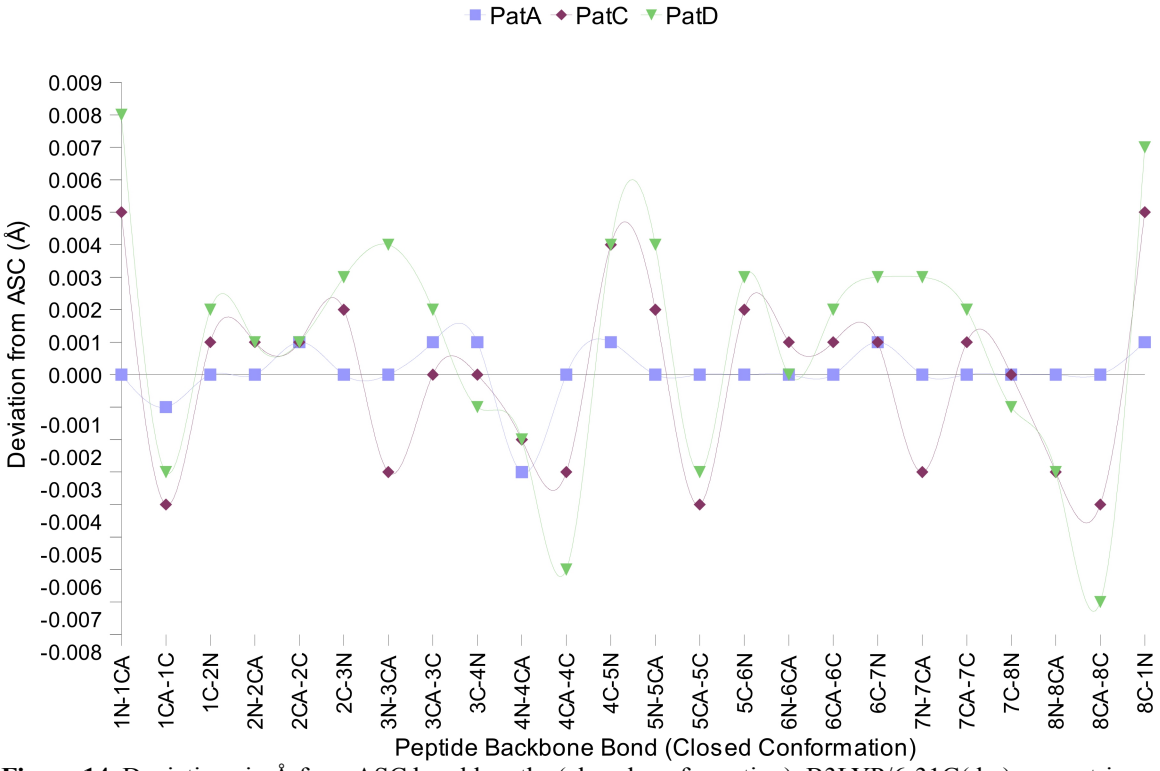


Figure 14. Deviations in Å from ASC bond lengths (closed conformation). B3LYP/6-31G(d,p) geometries.

The fact that PatD also displayed a relatively large deviation at residue 3, even though it is identical to ASC in this region, suggests that this effect must be induced by the substitutions in residues 1 and 5. Unlike PatA, the folding of the macrocycle into the closed conformation greatly increased the magnitude of the bond length deviations in PatC and PatD with values between 0.003 and 0.008 Å being found in residues throughout both structures (Figure 14). The deviations found in the closed conformer were generally larger for PatD than for PatC.

The open conformer bond angles (Figure 15) in PatA were less informative than the corresponding bond lengths for this molecule as deviations of similar magnitude occurred throughout the macrocycle. A clustering of slightly larger deviations was again seen around residue 4 but this was more localized than in the bond lengths. Again, the folding of PatA into its closed conformer seemed to lead to a redistribution of this small cluster into other regions of the macrocycle and a general reduction in the absolute values of the deviations. The bond angle deviations for PatC and PatD were found to be considerably larger than the corresponding values in PatA. In the open conformer the largest deviations for both peptides occurred at residue 1 (D-Phe) where the 1N-1CA-1C angle was reduced by ~1.5º relative to the D-Val residue in ASC. Deviations of ~1.0º were found for the 5N-5CA-5C angle (PatD = D-Ala, ASC = D-Val) and in angles throughout residues 2 and 8.

In the closed conformer the large deviations at 1N-1CA-1C were reduced almost to zero but the adjacent 8C-1N-1CA (negative) deviations were increased in magnitude from ~0.2º to 2.0º (PatC) and 3.5º (PatD). Similarly, the 5N-5CA-5C deviation was seen to be reduced in the closed conformation to approximately half of its 'open' value whilst the adjacent 4C-5N-5CA angle deviations changed from ~ +0.2º to -1.7º (PatC) and -2.5º (PatD).

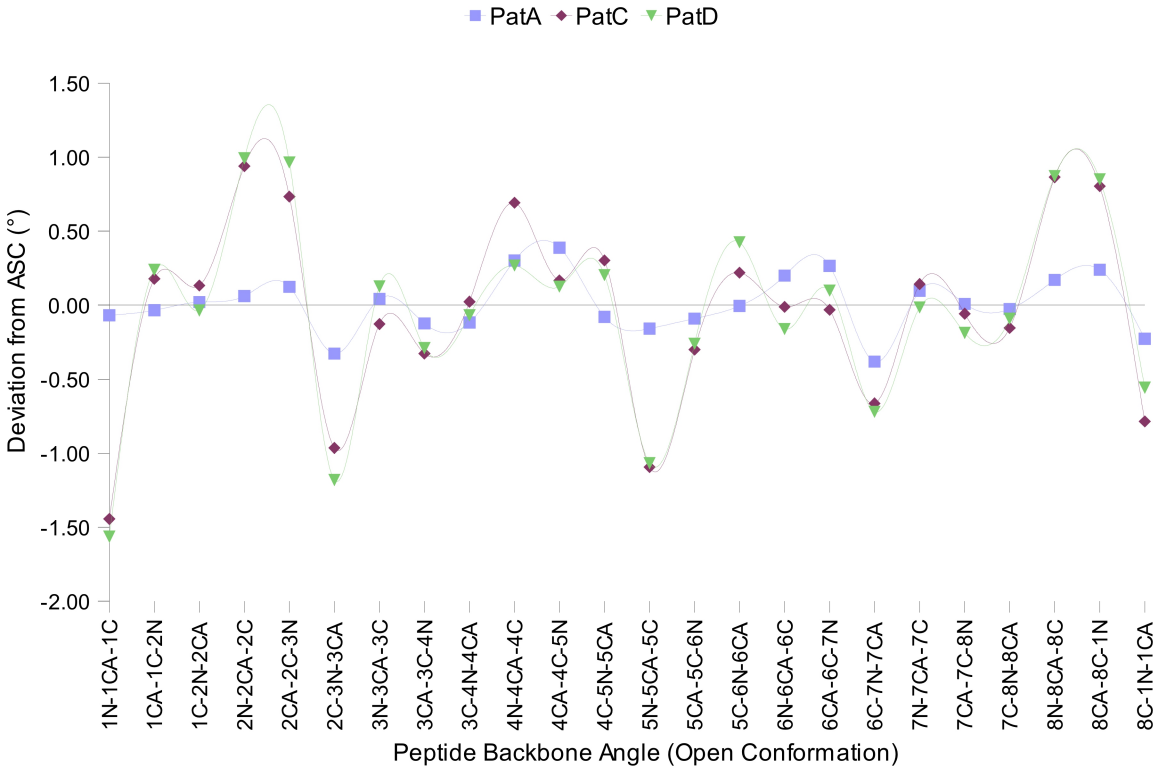


Figure 15. Deviations in º from ASC bond angles (open conformation). B3LYP/6-31G(d,p) geometries.

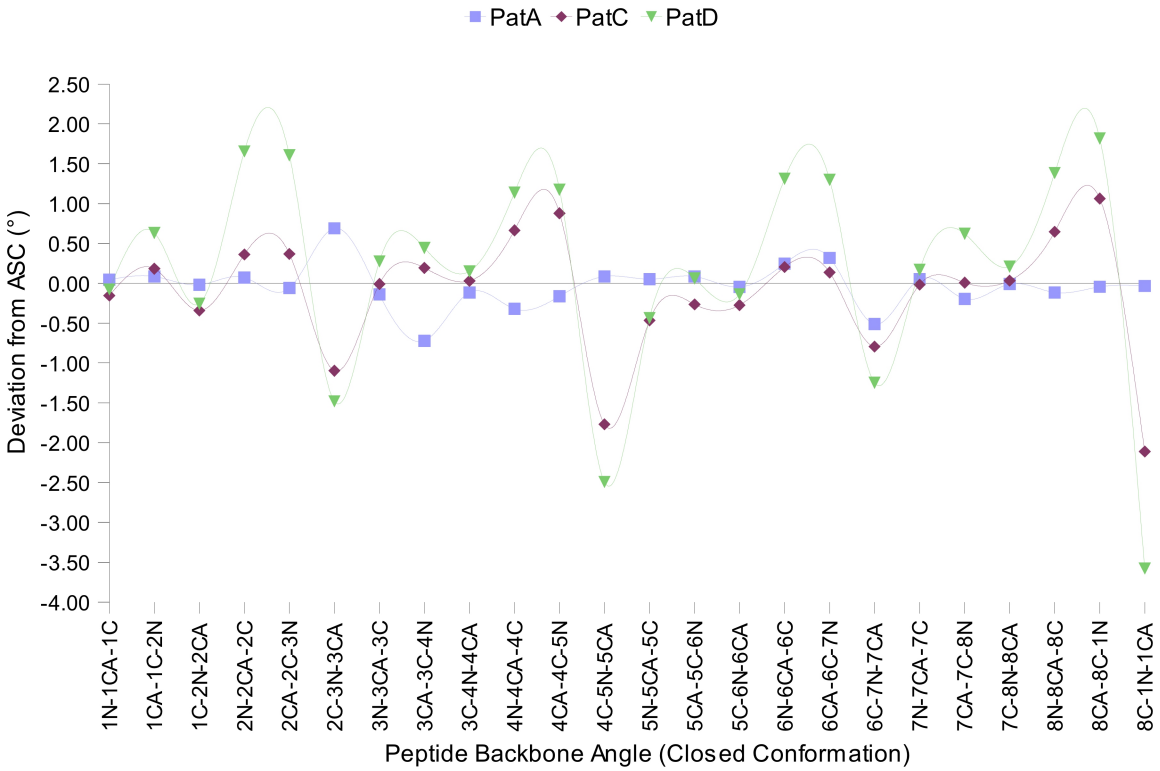


Figure 16. Deviations in º from ASC bond angles (closed conformation). B3LYP/6-31G(d,p) geometries.

The relative ease of deformation of dihedral angles (Figures 17 & 18) meant that this structural coordinate was expected to provide the most sensitive response to the various substitutions in the patellamide structures. The largest deviation in PatA (open conformer) was again found to occur close to the demethylated oxazoline ring. The dihedrals within residue 5 (between atoms 4C and 6N) showed a total deviation of 12º from those in ASC. Residue 1 (opposite residue 5 in the macrocycle) showed a similar total deviation of 7º. This was in agreement with the distribution of deviations in the bond lengths and angles in that it highlighted the fact that the removal of the methyl group from one of the oxazoline rings in ASC leads to significant perturbations in the structure at sites some distance from this position. On folding into the closed conformation the positions of maximum deviation in the PatA dihedrals were seen to move into the adjacent residues 3 and 4. Apart from the deviations in these residues, the closed conformation of PatA displayed lower overall deviation from the pattern of dihedrals in the corresponding ASC conformer.

The open conformers of PatC and PatD both had large dihedral deviations in residues 1 (~25º for PatC and ~30º for PatD) and 5 (~20º for PatC and ~13º for PatD). Overall, the deviations seen in these peptides were considerably larger than the corresponding PatA deviations. An exception to this was seen between residues 4 and 6 where PatD had similar dihedral structure to PatA whilst the deviations for PatC in this region were noticeably larger. The changes in the dihedral deviations of both PatC and PatD caused by moving to the closed conformation were more erratic when the two peptides were compared (Figure 18).

Large reductions in the deviations at residue 1 were noted for PatD. In PatC, one of the two deviations in this residue was seen to be reduced by ~50% relative to ASC whilst the other was increased by an insignificant amount (~2º). The similarity between residues 4-6 of PatD and PatA seen in the open conformation was lost on folding of the macrocycle, with differences in both the magnitudes and signs of the dihedral deviations being observed.

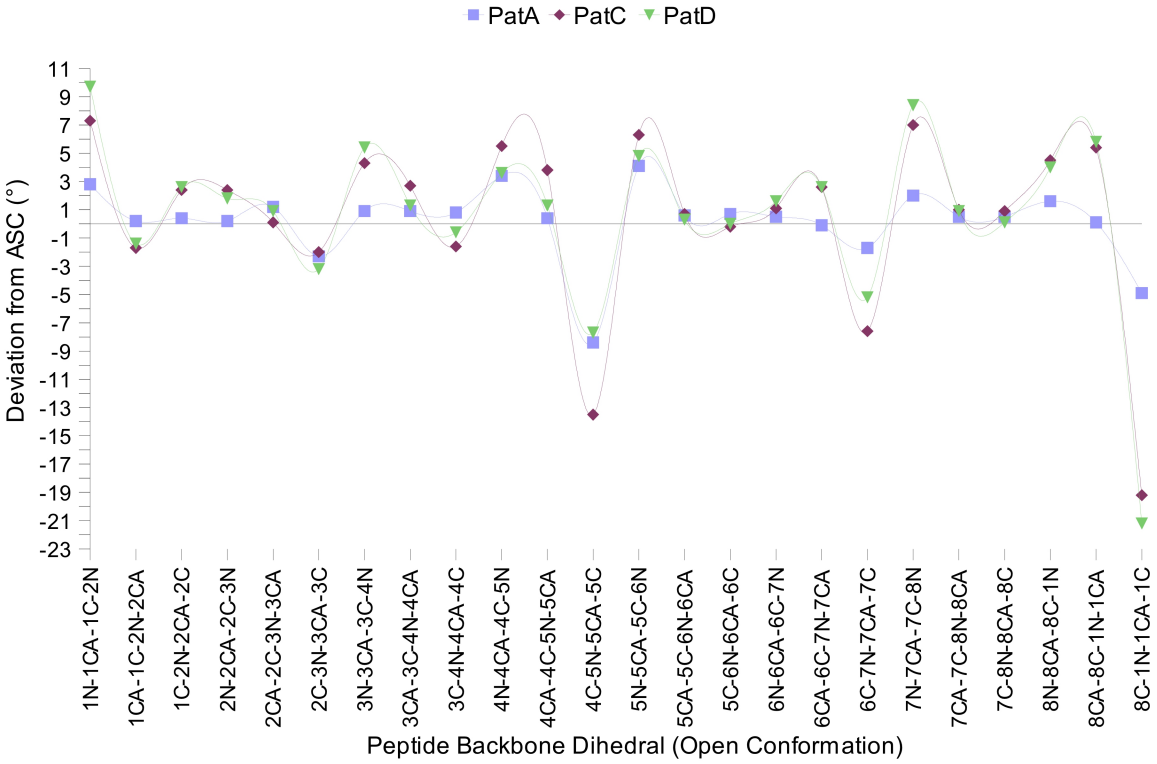


Figure 17. Deviations in º from ASC dihedrals (open conformation). B3LYP/6-31G(d,p) geometries.

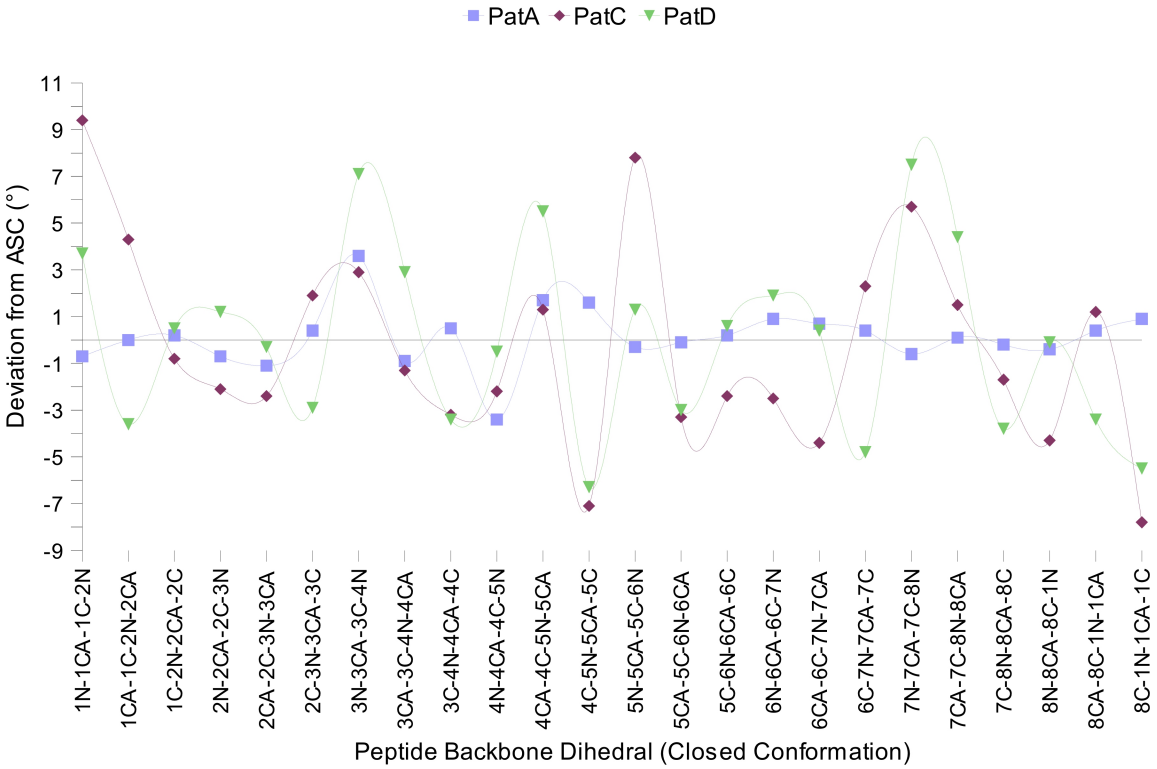


Figure 18. Deviations in º from ASC dihedrals (closed conformation). B3LYP/6-31G(d,p) geometries.

Discussion

From the relatively detailed comparisons of the patellamide macrocyclic geometries in the present work it can be seen that significant deviation from C_2 symmetry (as displayed by ASC) occur at both the local and global levels in the three substituted patellamides studied. Whilst the overall deviation between ASC and PatC/D is a somewhat trivial finding in light of the large differences between ASC and these peptides, the fact that PatA also displays noticeable deviations despite the fact that the backbone substitution pattern is identical to that of ASC is more surprising. Within the pairs ASC/PatA and PatC/PatD it can be seen that merely changing from CH_3 to H at positions outwith the macrocycle itself can have significant effects on the global geometric features (as well as the conformational energetics) of octapeptides of this type.

The critical importance of the oxazoline rings in the folding of the patellamides was previously highlighted in an extended molecular dynamics study performed within this group (see Appendix A - Milne *et al.*, 2002).[17] This study showed that the first step in the folding process was the rotation of the region of the macrocycle containing these features and that this subsequently allowed the remainder of the patellamide structure to undergo the changes necessary to produce the closed conformation. The detailed differences in the equilibrium geometrical coordinates which were observed in the present work support the importance of these structural subunits and give some indication of how local substitutions to the ASC structure could lead to changes in the energetics of the folding process itself. Changes in equilibrium bond lengths, angles and dihedrals would be accompanied by corresponding changes in the force constants governing their dynamic properties. Thus, the equilibrium changes reported here can reasonably be taken to indicate differences in the dynamic properties of the patellamides and consequently of their energetic profiles in conformational space.

Whilst it is difficult to quantitatively summarize the global geometrical effects of the various substitutions, these effects are strongly reflected in the energetic data obtained in the present study. The energies obtained from the B3LYP/6-31G(d,p) calculations relate only to the equilibrium geometries of the open and closed conformers and consequently cannot provide any information on the energetics of the folding process itself. Despite this

shortcoming, if it is assumed that thermodynamic equilibrium between the two conformations exists then these global energy changes can be taken to be indicative of the conformational preference in the patellamides studied. Taking such a view, the experimentally derived trend favouring folding in the more highly substituted patellamides is predicted well from these calculations. The fact that all four peptides display positive energy changes seems to contradict this trend until the effects of solvation (or the crystal environment of solid state structures) are taken into account. The gas phase energy differences observed for PatC/D are of small enough magnitude to allow the reasonable inference that solvation may be responsible for altering the conformational preferences in these peptides. The experimental observation that PatA assumes the open conformation in (polar) MeOH [17,19] and the closed conformation in (non-polar) $CDCl_3$ [17] lends more weight to this supposition.

Because the magnitude of the positive gas phase energy difference is approximately three to four times larger in PatA than in PatC/D and neither PatC or PatD have been observed in the open conformation, if solvation reduces the energy difference in PatA to a point where the conformational preference is readily modified by changes in the polarity of the solvent then the solvation energy can quite reasonable be expected to reduce the conformational energy difference in the more highly substituted patellamides to the point where the open conformation is almost totally unpopulated. This insensitivity to solvent polarity is found with PatC with the closed conformation being adopted in both MeOH [17] and $CDCl_3$ [15]. If PatA is correctly placed (energetically) to become susceptible to solvent polarity but PatC/D are not, it might be supposed that ASC could also be insensitive to this type of environmental effect, as it has the largest conformational energy differences of all the compounds studied here. This is also supported by experimental observation, with the open conformation being adopted in $CDCl_3$ [13] contrary to all of the other members of this family.

The comparisons with the energies obtained from the molecular mechanics calculations with the MM2* and AMBER* force fields have highlighted an important deficiency with respect to the use of this approach to study the conformational behaviour of the patellamides (and possibly other similar cyclic peptides). MM2* performed reasonably

well in placing the patellamides in the correct order with respect to the conformational trend but failed to predict the correct (B3LYP/6-31G(d,p)) gas phase conformational energy changes. AMBER* (which is commonly used in studies on peptides and proteins) failed to predict either the correct trend in the patellamides studied or the correct energy changes connecting the open and closed conformers. This is an important finding both for future studies on cyclic peptides of this type and for the interpretation of data already existing in the literature as incorrect assumptions could quite easily be drawn from the results of this type of calculation if this source of error was to remain unnoticed.

Chapter 3

Discussion of Non-Covalent Interactions in the Patellamides

Charge-transfer involving thiazole units

Because of the close proximity and almost parallel distribution of the thiazole ring subunits observed in the closed conformers of the patellamides it has been suggested that charge-transfer (CT) interactions involving the π-systems of these units (transannular π-stacking) could occur in these molecules.[24,25] Such an effect, if it exists in these peptides, could have a significant effect on the conformational energetics of the patellamides by lowering the energy of the closed conformation relative to that of the open conformer in which the thiazole units are as far from each other as is possible without breaking the macrocyclic ring.

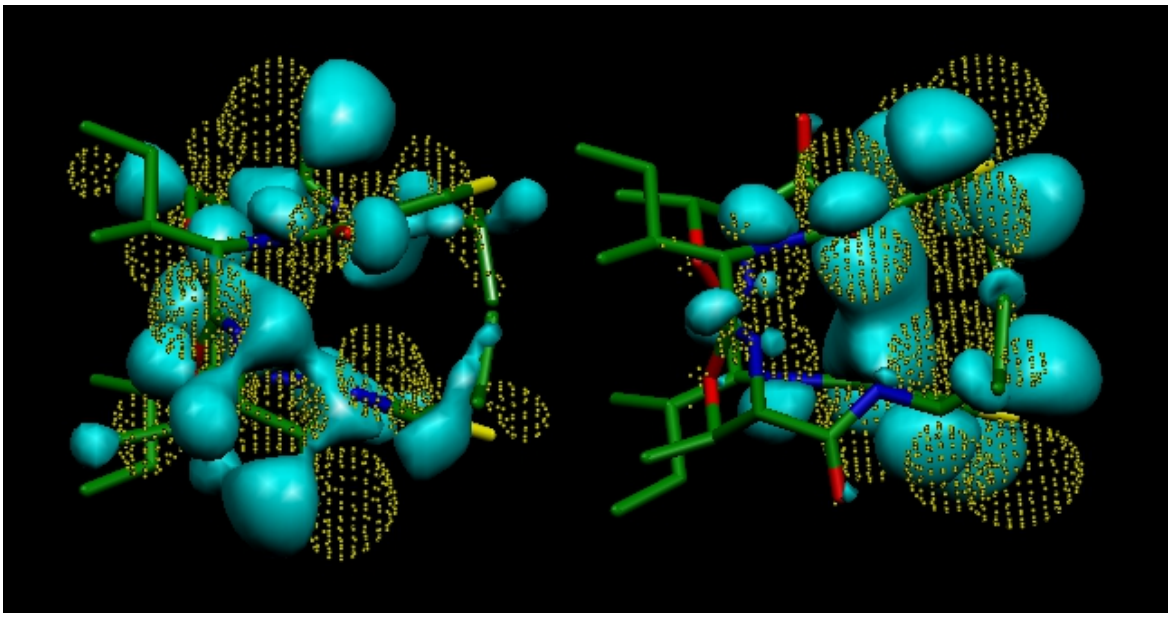


Figure 19. HOMO (left) and LUMO (right) of ascidiacyclamide closed conformer (surfaces correspond to 0.01 electron cutoff). To increase clarity hydrogens have been omitted and the positive regions of the orbitals have been shown as dot surfaces.

Visual inspection of the Kohn-Sham molecular orbitals of the patellamides (taken here to be approximate indicators of the spatial distribution of the relevant donor and acceptor regions of the total change density) shows that the LUMO, which would act as the charge-acceptor in such an interaction, is mainly localized on the two thiazole ring-containing regions of the molecules (ascidiacyclamide is given as an example in Figure 19). This adds weight to the possibility of an energy-lowering transannular charge-transfer in the closed conformation as bonding interactions between the thiazoles would be expected to result from population of this orbital. The HOMO (charge donating region) is seen to

reside mainly around the oxazoline ring-containing regions and therefore no favourable interaction between the thiazoles would be lost by the depopulation of this orbital.

In order to find out whether alterations in this transannular separation might occur in the variously substituted patellamide structures in the present study, 'dummy' atoms were placed in the center of the thiazole rings of ASC and PatA/C/D and the distance between these points measured. The results of these measurements are shown in Table 13 and display large reductions (~0.3 Å) on going from ASC/PatA to PatC/D.

Patellamide	Transannular separation (Å)
ASC (**1**)	4.34
PatA (**2**)	4.33
PatC (**4**)	4.09
PatD (**5**)	4.04

Table 13. Transannular separation of thiazole rings in patellamide closed conformers (B3LYP/6-31G(d,p) geometries). Values correspond to measurements made between centers of thiazole rings.

If the assumption is made (for the purposes of this discussion) that this effect is not simply due to the constraints placed on this coordinate by the macrocycle itself and that CT interactions are solely responsible, explanation of the apparently increased attraction between the thiazole units in PatC and PatD requires that a suitable alteration in the properties of this region of the two peptides, relative to ASC/PatA, should exist. Such an alteration exists in the form of the substitution of the two D-Val residues in ASC/PatA by D-Ala and D-Phe residues in PatC/D. Aromatic amino acid residues are known for their charge-donating properties[60] and the fact that the D-Phe residue in PatC/D is located immediately adjacent to one of the thiazole units in these peptides suggests that redistribution of charge between D-Phe and thiazole might be expected to occur through the bonds connecting these moieties. That the side 'chain' adjacent to the other thiazole unit in the structure is a simple (charge-withdrawing) methyl group has the consequence that the two thiazole-containing regions might be expected to display differing CT properties to one another in PatC/D.

The effects of differences in two chemical species on the net CT occurring between them can be shown using arguments from density functional theory. The degree and direction of

the net ground state CT between two chemical entities A and B is given (to second order) by[61-63]

$$\Delta N = \frac{1}{2}\frac{(\mu_B - \mu_A)}{(\eta_A + \eta_B)} \tag{3.1}$$

Here, μ_x is the electronic chemical potential which is defined as the functional dependance (designated by "δ" in equation 3.2) of the total energy of species x, E_x, with respect to changes in the charge density, ρ_x, when the external potential, $V_x(\vec{r})$, is held constant; this is equivalent to the first derivative of E_x with respect to the electron number, N_x:[61,62,64,65]

$$\mu_x = \left(\frac{\delta E_x}{\delta \rho_x}\right)_{V_x(\vec{r})} = \left(\frac{\partial E_x}{\partial N_x}\right)_{V_x(\vec{r})} \tag{3.2}$$

Physically, μ_x can be interpreted as the capacity of species x to donate electron density and consequently the net flow of charge between two reactants will be from high to low μ. The quantity η_x is the chemical 'hardness' of species x and is a measure of the resistance to redistribution of ρ_x resulting from variations in N_x. Thus, the values of η determine the degree to which charge is redistributed between two species. The chemical hardness is defined as the first derivative of μ_x:[61,62,65,66]

$$\eta_x = \left(\frac{\partial \mu_x}{\partial N_x}\right)_{V_x(\vec{r})} = \left(\frac{\partial^2 E_x}{\partial N_x^2}\right)_{V_x(\vec{r})} \tag{3.3}$$

The charge transfer energy, ΔE^{CT}, associated with the redistribution of a quantity of charge, ΔN, between systems A and B may be obtained by using the expression[61-63]

$$\Delta E^{CT} = -\frac{1}{4}\frac{(\mu_B - \mu_A)^2}{(\eta_A + \eta_B)} \quad . \tag{3.4}$$

This serves to illustrate the point that the presence of charge-donating/-withdrawing groups adjacent to the thiazole rings in PatC and PatD can be expected to have an effect on the CT properties of these units by inducing changes in the values of μ (and consequently in the values of η). Whereas in ASC both of these units are identical, and would therefore display zero net CT, the intramolecular difference between the thiazole units in PatC/D should lead to a non-zero value for ΔE^{CT} and therefore increased overlap (reduced transannular separation) of the two thiazoles in these patellamides would be favoured. The corresponding net transannular CT in PatA would be expected to show only minimal deviation from that of ASC and this is supported by the difference of only 0.01 Å in the transannular distances of the two peptides (~3% of the PatC/ASC and PatD/ASC differences).

The argument given above is clearly physically incomplete in that it assumes that charge transfer is the factor governing the transannular separation of the thiazole units when the effects of the peptide backbone cannot in reality be neglected. If the converse view is taken that the transannular separation in the closed conformation is completely controlled by the steric properties of the peptide macrocycle, the energy lowering due to transannular CT (assuming non-zero net CT between fragments) would still be expected to be greater for unsymmetrical patellamides because of increased overlap of the relevant donor and acceptor densities.

Hydrogen bonding

A similar argument to that given above for intramolecular charge-transfer can be applied to hydrogen bond stabilization of the closed conformation of the patellamides. Four hydrogen bonding interactions are possible in the closed conformation, as shown in Figure 20 (in the open conformation the arrangement of hydrogen bond donors and acceptors is such that no interactions are possible). The charge acceptors in these interactions are the four amide hydrogens spaced equally around the macrocycle. Charge donation is from the two amide carbonyl oxygens adjacent to the thiazole rings and from the two oxygens incorporated into the oxazoline ring units.

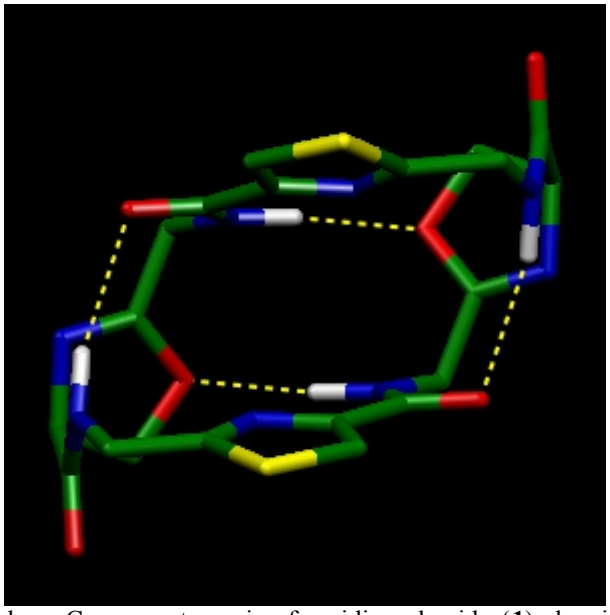


Figure 20. View along C_2 symmetry axis of ascidiacyclamide (**1**) showing hydrogen bond "cage" (yellow dotted lines) present in the closed conformation of the patellamides. Side chains and backbone hydrogens have been removed to improve clarity.

	Carbonyl donor oxygen		Oxazoline donor oxygen	
Patellamide	1NH-6O	5NH-2O	3NH-7O	7NH-3O
ASC (**1**)	2.406	2.406	2.491	2.491
PatA (**2**)	-0.017	-0.010	-0.025	-0.106
PatC (**4**)	-0.131	-0.167	-0.170	-0.107
PatD (**5**)	-0.043	-0.075	-0.015	-0.085

Table 14. Hydrogen bonding in patellamide closed conformation (B3LYP/6-31G(d,p) geometries). Distances are shown explicitly for ASC and deviations from these values are shown for PatA/C/D. Measurements in Å.

From Figure 20 the H-bonds containing the carbonyl oxygen donors can be seen to be involved in holding together the two sides of the pseudo type-II β-turns formed around the oxazoline-containing segments of the macrocycle. The remaining hydrogen bonds serve to restrict the torsional rotation of the oxazoline rings which is known to be play a crucial part in the overall folding process.[17] The distances corresponding to these hydrogen bonds in the four patellamides studied in the present work are shown in Table 14.

The symmetry of ASC is evident even in these non-bonded interactions as is the breaking of this symmetry in PatA/C/D. In PatA the largest deviation from the ASC distances occurs in H-bond 7NH-3O which involves the oxygen atom in the demethylated oxazoline ring. In PatD the deviations are of similar magnitude to those for PatA and again the largest deviation is in 7NH-3O. In PatC this coordinate is also much shorter than in ASC,

however this is the smallest of all the deviations for this patellamide with the other H-bonds showing much larger deviations from the ASC distances than either PatA or PatD. Determining whether these deviations are a result of changes in the donor/acceptor properties of the relevant regions of the molecules due to the various substitutions or simply reflect the net effect of differences in the bonded coordinates is again impossible without explicitly tailored calculations. However, lowering of the total energy can be inferred from the decreases in the hydrogen bond lengths corresponding to strengthening of these interactions (assuming that these interactions do not become so strong as to cause deviations from the equilibrium bond geometry which would lead to increases in energy) and substitutional symmetry can be seen to mediate this effect, regardless of whether the effect is direct or indirect. The fact that these hydrogen bonds are known from molecular dynamics simulations[17] to form, break and reform several times during the folding process for these molecules suggests that changes in the intrinsic strengths of these interactions due to substitution should have an effect on the stability of the intermediate conformations involved in the transition between the open and closed conformers with direct consequences for the conformational preferences displayed by the different peptides.

Effect of conformation and substitution on ground-state global charge-transfer properties of the patellamides

Parr *et al*, 1999,[67] showed that setting the chemical potential of the environment surrounding a molecule to zero will lead to maximal CT, ΔN_{MAX}, from the surroundings to the species in question. Under these conditions the maximum energy lowering possible by charge transfer, ΔE_{MAX}^{CT}, will be obtained, giving a measure of the intrinsic charge accepting capability (electrophilicity) of the system in the ground state. The authors suggested this as an approximation to the binding environment of an electrophilic ligand interacting with a surface or a large biomolecular assembly such as a protein or a DNA coil.[67] The quantities μ and η (equations 3.2 and 3.3, respectively) may be combined to give a second-order expression for ΔE_{MAX}^{CT}:[67]

$$\Delta E_{MAX}^{CT} = -\left(\frac{\partial E}{\partial N}\right)_{v(\vec{r})} \Delta N_{MAX} + \frac{1}{2}\left(\frac{\partial^2 E}{\partial N^2}\right)_{v(\vec{r})} \Delta N_{MAX}^2$$
$$= \mu \Delta N_{MAX} + \frac{1}{2}\eta \Delta N_{MAX}^2 \tag{3.5}$$

Here, ΔN_{MAX} is given by [67]

$$\Delta N_{MAX} = -\frac{\mu}{\eta} \quad . \tag{3.6}$$

In practice, the calculation of μ and η is simplified by the introduction of the finite difference approximations

$$\mu = \frac{1}{2}(I + A) \tag{3.7}$$

and

$$\eta = (I - A) \tag{3.8}$$

where I is the first ionization potential and A is the first electron affinity of the system in question. A good estimate estimate of I can be obtained from the results of quantum chemical calculations by the application of Koopmans' theorem:[40]

$$I \approx |\epsilon_{HOMO}| \tag{3.9}$$

where ϵ_{HOMO} and ϵ_{LUMO} are the energies of the highest occupied and lowest unoccupied molecular orbitals, respectively. A similar approach can be used in estimating the magnitude of A

$$A \approx |\epsilon_{LUMO}| \tag{3.10}$$

It should be noted, however, that this approach can be unreliable as the eigenvalues corresponding to virtual molecular orbitals are heavily dependant on the basis set used. This approach to estimating A was eventually used because of the size of the systems involved and the difficulties involved in performing open-shell calculations on them.

Substituting μ and η in equations 3.5 and 3.6 with their finite difference approximations provides the operationally useful expressions

$$\Delta E_{MAX}^{CT} = -\frac{I+A}{2}\Delta N_{MAX} + \frac{I-A}{2}(\Delta N_{MAX})^2 \tag{3.11}$$

$$\Delta N_{MAX} = \frac{(I+A)}{2(I-A)} \tag{3.12}$$

Using the values of ϵ_{HOMO} and ϵ_{LUMO} obtained from the B3LYP/6-31G(d,p) calculations performed earlier, the maximal global charge transfer energies, ΔE_{MAX}^{CT}, of the lowest energy rotamers (in both the open and closed conformations) of ASC, PatA, PatC and PatD were calculated using equations 3.11 and 3.12. The results of these calculations are shown in Table 15.

Peptide	Conformer	I	A	μ	η	ΔN_{MAX}	ΔE_{MAX}^{CT}
ASC (**1**)	open	6.73	1.31	-4.02	5.42	0.74	-1.49
	closed	6.52	1.32	-3.92	5.20	0.75	-1.47
PatA (**2**)	open	6.75	1.31	-4.03	5.43	0.74	-1.50
	closed	6.54	1.34	-3.94	5.19	0.76	-1.49
PatC (**4**)	open	6.65	1.30	-3.97	5.36	0.74	-1.47
	closed	6.46	1.39	-3.93	5.07	0.77	-1.52
PatD (**5**)	open	6.65	1.32	-3.98	5.32	0.75	-1.49
	closed	6.45	1.39	-3.92	5.07	0.77	-1.52

Table 15. Properties relating to partial charge transfer in the patellamides. I and A (eV, obtained from results of B3LYP/6-31G(d,p) calculations), DFT reactivity parameters μ and η, maximum charge transfer, ΔN_{MAX} (electrons), and maximum energy lowering associated with charge transfer from surroundings, ΔE_{MAX}^{CT} (eV).

Folding into the closed conformation leads to an increase in the quantity of charge (ΔN_{MAX}) that can be accepted by all four patellamides. From consideration of this parameter alone it would seem that the closed conformer should be the most electrophilic in each of the molecules studied. The energy changes relating to this charge-transfer, however, show that this is the case only for the two phenylalanine-containing peptides PatC and PatD. ASC and (to a lesser extent) PatA both gain the largest energy lowering due to CT when they are in the open conformation.

This is an interesting result as the difference in the conformation-dependent maximal ground state charge-transfer energy in the gas phase, $\Delta\Delta E_{MAX}^{CT}$, is much more in line with the experimentally observed conformational preferences displayed by these molecules in solution than the B3LYP/6-31G(d,p) conformational energies, ΔE^{DFT}. Both of these energy changes are shown in Table 16 (values in kJ mol^{-1}). The positive energy change for ASC and the negative changes for PatC/D agree with the preference for the open conformation in the former and the closed conformation in the latter two, whereas PatA is on the borderline with only a slight preference for the open conformation. The effect of adding $\Delta\Delta E_{MAX}^{CT}$ to ΔE^{DFT} to obtain gas phase conformational energies corrected to

include the effects of charge-transfer, ΔE^{TOTAL}, is also shown in Table 16.

Peptide	$\Delta\Delta E^{CT}_{MAX}$	ΔE^{DFT}	ΔE^{TOTAL}
ASC (**1**)	1.7	21.4	23.1
PatA (**2**)	0.04	16.4	16.4
PatC (**4**)	-4.5	5.8	1.3
PatD (**5**)	-2.4	3.6	1.2

Table 16. Effect of conformational change on global maximum ground state charge-transfer energies of the patellamides $(\Delta\Delta E^{CT}_{MAX})$. ΔE^{TOTAL} is the result of including $\Delta\Delta E^{CT}_{MAX}$ as a correction to ΔE^{DFT} (obtained from B3LYP/6-31G(d,p) calculations). $\Delta\Delta E^{CT}_{MAX}$ for PatA is shown to two decimal places in order that its non-zero value can be seen. All values in kJ mol^{-1}.

The observation that favourable charge-transfer leads to a reduction in the magnitude of the positive energy change associated with folding of PatC/D in the gas phase may help to rationalize the unfolding of these patellamides that occurs on interaction with the positively charged Cu^{2+} ion in solution.[20] Whilst deprotonation of one of the amide nitrogens involved in the hydrogen bonds is apparently crucial for this unfolding (other dipositive metal ions such as Zn^{2+} do not effect this deprotonation and no conformational change as great as that for Cu^{2+} are observed) the charge-withdrawing properties of Cu^{2+} could be expected to reduce the effect of $\Delta\Delta E^{CT}$ on the conformational preferences of PatC/D.

Such a charge-withdrawing mechanism could also have a destabilizing effect on the intramolecular transannular CT discussed above. Empirical evidence which seems to support this comes from circular dichroism and restrained molecular dynamics studies on the complexation of Zn^{2+} to PatC.[20] In this study the binding of Zn^{2+} resulted in the partial unfolding of the patellamide and an increase in the transannular separation of the thiazole ring units, as would be expected if the postulated π-stacking interaction between these units was to be diminished by interaction with a strong charge-acceptor.

Effect of varying transannular separation on $\Delta\Delta E^{CT}_{MAX}$ for PatD

A possible explanation for the differences in the values of $\Delta\Delta E^{CT}_{MAX}$ for the patellamides studied here comes from the qualitative discussion of transannular CT given previously. If,

in the closed conformation, non-zero net CT can occur within the molecule (and the CT capacity is not saturated by the molecule itself) then a further lowering of energy might be expected on addition of extra charge from an external source. If such a mechanism operates within in the patellamides, increasing/optimizing the overlap between portions of the molecule likely to take part in such a stabilizing interaction might be expected to increase the charge-transfer capability of the system as a whole.

The effect on ΔE_{MAX}^{CT} of variations in the transannular separation of the thiazole rings was investigated for PatD as this coordinate was much shorter in the XRD structure[25] than in the relaxed gas-phase B3LYP/6-31G(d,p) structure obtained in the present work. A single-point B3LYP/6-31G(d,p) calculation was performed at the XRD geometry using the twin-processor workstation at Aberdeen and ΔE_{MAX}^{CT} calculated as before. The results of this calculation along with the corresponding gas phase values are shown in Table 17.

PatD (**5**)	Transannular separation (Å)	μ	η	ΔN_{MAX}	ΔE_{MAX}^{CT}
Closed (DFT)	4.04	-3.92	5.07	0.77	-1.52
Closed (XRD)	3.62	-4.02	4.91	0.82	-1.64

Table 17. Change in global charge-transfer properties on varying transannular separation of thiazole rings in patellamide D. Units as in Table 15.

The increase in the magnitude of ΔE_{MAX}^{CT} observed on reducing the transannular separation was -0.12 eV (-11.6 kJ mol^{-1}) corresponding to an increase in ΔN_{MAX} of 0.05 electron units. Unfortunately, since the XRD structure was not locally optimized at the B3LYP/6-31G(d,p) level, comparison with the open conformer (to obtain a new corrected conformational energy change) was not possible. Whilst it is true that this also applies to the comparison between the XRD and B3LYP/6-31G(d,p) closed conformers performed here, this result seems to support the idea that further decreasing the transannular separation between the thiazole units (relative to the relaxed B3LYP/6-31G(d,p) closed conformer geometry) should improve the overlap of the relevant regions of the charge density and consequently lead to a larger energy-lowering.

Conformation-dependent changes in the dipole moments of the patellamides

The sensitivity to the polarity of the solvent environment shown by the conformational preference in the patellamides[17,68] suggested that the molecular dipole moment was a property that might be important in these peptides. Dipole moments for the four patellamides studied here were obtained at the end points of the B3LYP/6-31G(d,p) geometry optimizations and are shown in Table 18.

Peptide	"open"	"closed"	Δ
ASC (**1**)	3.04	1.79	-1.25 (-41.2%)
PatA (**2**)	3.24	1.54	-1.70 (-52.5%)
PatC (**4**)	2.92	2.16	-0.72 (-26.0%)
PatD (**5**)	2.80	1.79	-1.01 (-36.1%)

Table 18. Conformational dependance of patellamide dipole moments calculated at the B3LYP/6-31G(d,p) level. Values in Debye units.

Inspection of the values in Table 18 shows that in all four patellamides the dipole moment is reduced on going from the open to the closed conformations. In the open conformation the magnitudes of the dipole moments are seen to decrease in the phenylalanine-substituted patellamides relative to ASC and PatA, with PatA having the largest moment of all four peptides. In the closed conformation PatC has the largest dipole moment (2.16 Debye) and PatA the smallest (1.54 Debye) with both PatD and ASC having the same value of 1.79 Debye. The change in the magnitude of the dipole moment on folding of the macrocycle was largest for PatA (-52.5%) and smallest for PatC (-26.0%).

Both PatC and PatD showed smaller changes than the more symmetrical ASC and PatA. This is not the result that might be expected from consideration of the previous observation that polar solvents lead to a favouring of the closed conformation, suggesting that the gas phase dipoles given here may undergo considerable alteration upon solvation. Alternatively, it may be that the dipole moment does not play a dominant part in determining the solution conformation of these molecules.

An interesting result was obtained when the orientation of the dipole moments with respect to the patellamide structures was plotted graphically (Figures 21-24). Here, it can

be seen that in addition to undergoing a reduction in magnitude the dipoles are re-oriented through 180º. In ASC the dipole (represented by a red arrow, the head of which corresponds to the negative pole) is aligned along the C_2 symmetry axis of the molecule. The effects of the desymmetrization in the other patellamides are evident in the deviation of the dipole arrow from the position of this axis in ASC.

Inversion of the dipole by the folding of the patellamides may have consequences for any biological activity shown by these peptides in that reversal of the direction and alteration in the magnitude of the dipole moment, if it were to occur whilst the molecule was bound to an active site or surface, could act to mediate the strength and type of the interaction and alter the properties of the complex. In the relatively simple (hypothetical) case of interaction with a negatively charged biological membrane surface (it is assumed that no binding site specific for the patellamide structure is involved), the patellamide would be expected to orient itself so that the positive end of the dipole came into contact with the membrane surface. Whilst this arrangement would be stable, if the patellamide were to undergo conformational change the dipole would be reoriented so that the negative end was now pointed at the negatively charged membrane and consequently the interaction would become unfavourable, leading (in the extreme case) to dissociation of the patellamide from the surface.

Such a model seems even more plausible when the effects of charge-transfer are included as the acceptance of charge from the membrane could have significant effects on the conformational preference of the patellamide involved. Whilst this hypothetical model neglects effects such as solvation/desolvation during the different stages of the interaction with the membrane and changes in the dipole moment caused by the acceptance of charge, as a first approximation it is quite intriguing especially as previous authors working on these compounds have pointed out the possibility that the patellamides could have an electron-transfer rôle in nature and ascidiacyclamide has been shown to interact strongly with model lipid bilayers in studies using differential scanning calorimetry.[13,24]

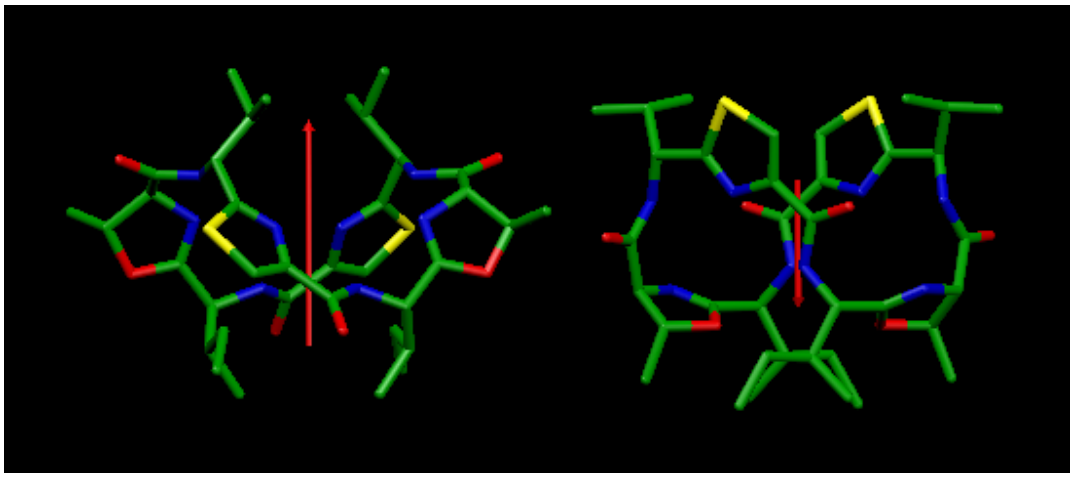


Figure 21. Orientation and relative magnitude of the dipole moment in ascidiacyclamide open and closed conformations. Hydrogens omitted to improve clarity.

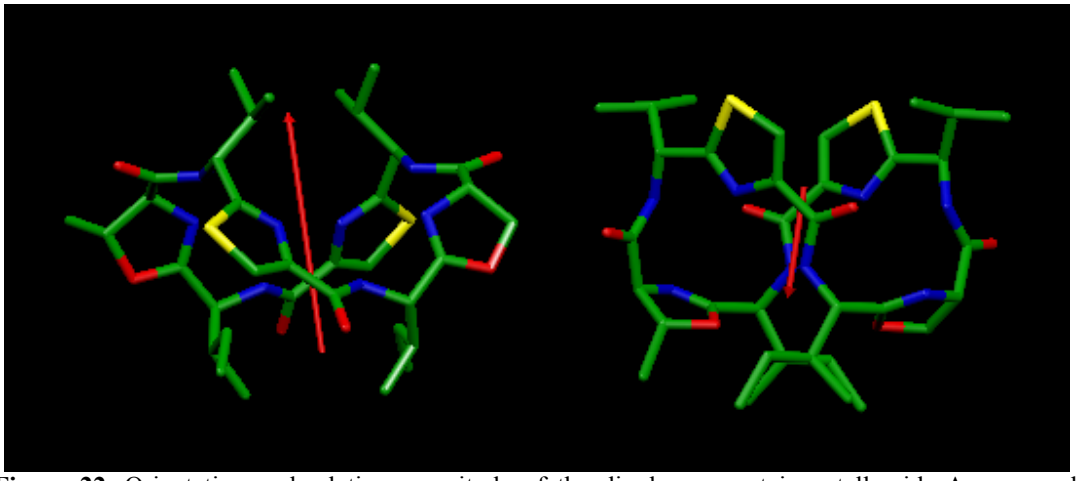


Figure 22. Orientation and relative magnitude of the dipole moment in patellamide A open and closed conformations. Hydrogens omitted to improve clarity.

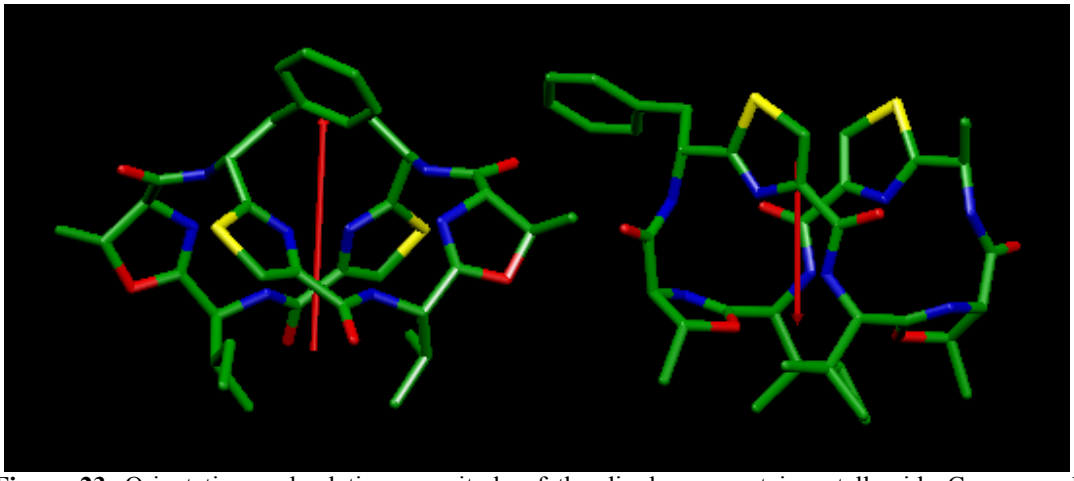


Figure 23. Orientation and relative magnitude of the dipole moment in patellamide C open and closed conformations. Hydrogens omitted to improve clarity.

Figure 24. Orientation and relative magnitude of the dipole moment in patellamide D open and closed conformations. Hydrogens omitted to improve clarity.

Part 2

Introduction

Metal binding properties of the patellamides

Metal binding by the patellamides has been studied by a number of experimental methods including X-ray crystallography, electron paramagnetic resonance (EPR), circular dichroism (CD), mass spectrometry (MS) and nuclear magnetic resonance (NMR).[18,19,68-71] These studies have focussed not only on the naturally derived compounds but also on a variety of synthetic analogues, the latter being used for comparison with the natural products in an attempt to better understand the nature of the metal-peptide interactions.[72,73] As it is the elucidation of details of the natural systems that is the focus of the present work, these analogues will not be considered in any depth here. The results of the metal studies on the natural patellamides are summarized below. All the metals chosen for these studies are divalent species from row 4 of the periodic table that are commonly found in biological systems and discussions of their behaviour are to be found in any modern text on (bio-) inorganic chemistry.[74-77]

Ca^{2+}

The interactions of both ascidiacyclamide and patellamide D with this metal have been studied.[72] It was found that both did form complexes with Ca^{2+}, with a number of species being characterised by NMR. Binding constants estimated from the NMR data were found to be particularly low (logK = 2.9±0.4 and 2.8±0.4 for ascidiacyclamide and patellamide D 1:1 metal - ligand complexes, respectively) when compared to those known for other cyclic peptides (logK = 4-7). CD spectra obtained during titration experiments in which the metal was added slowly to a peptide solution showed negligible differences indicating that the metals did not induce any change in the global peptide structure. In combination with the rate constants, these data were interpreted as meaning that only weak, non-specific binding of Ca^{2+} was occurring.

Zn^{2+}

Zinc complexation has been studied for ascidiacyclamide[18] and patellamides A and C[20,68]. Zn^{2+} was found to bind to patellamide C, inducing partial unfolding of the macrocycle from its prefered[15] closed conformation, as indicated by changes in the CD spectrum and by refinement of nOe data using restrained molecular dynamics.[20] No change was observed in the CD spectra of ascidiacyclamide[18] or patellamide A,[20] but this was to be

expected as both compounds show a preference for the open conformation in polar solvents such as the methanol used in these experiments.[68,78] MS data showed that both ascidiacyclamide and patellamide A[20] lost amidic protons in the zinc-bound state whilst patellamide C displayed no deprotonation. Significantly, the mass spectra of the patellamide A/Zn^{2+} species revealed the presence of a number of adducts in which more than two Zn^{2+} ions were detected.[71] This was taken as an indication that the binding of the metal to the patellamide was occurring in a non-specific manner (relative to the binding of Cu^{2+} where a stoichiometric ratio of 1:2 (patellamide:metal) was observed - see next section). Binding constants for patellamide A of $K_1 = 3 \times 10^4$ [19,71] have been calculated using data from titration experiments monitored by CD, with K_2=1000 [71] and 16 [19] also being found. Values of 1.8×10^4 for K_1 and 806 for K_2 in the patellamide C system were reported.[71]

The combined data suggests that whilst Zn^{2+} complexed more strongly to the patellamides than Ca^{2+}, the interaction is again relatively non-specific. Unlike for Ca^{2+}, Zn^{2+} containing systems appear able to effect deprotonation of amide nitrogens in the patellamides, thus providing a better charge-environment for the accommodation of the metal. The studies with the patellamides showed, however, that if these peptides are not already in the correct (open) conformation to allow sufficient access to those atoms then deprotonation does not occur.

Cu^{2+}

Ascidiacyclamide[69] and patellamides A,[71] C [71] and D [70] have all been the subject of investigations involving this metal. That all four patellamides form di-copper complexes in solution has been shown by ESR,[69,70] CD,[19,70,71] MS [70,71] and magnetic susceptibility measurements[69]. A feature of the patellamide-Cu^{2+} interaction that differs completely from that of either Ca^{2+} or Zn^{2+} is the complexation by patellamides C and D. Titrations of each with Cu^{2+} solutions gives large changes in the CD spectra of the peptides indicating that the metal is effecting a conformational change in the ligand. [19,71] In the study of patellamide C,[71] two peaks were monitored at 210 and 250 nm and on addition of one equivalent of Cu^{2+} the 250 nm positive maximum was seen to decrease whilst the 210 nm signal increased in intensity. Negligible changes were associated with addition of a second

equivalent of the metal. The signal at 250 nm was attributed to the β-turn present in the closed conformation and its decrease was taken to represent the peptide changing conformation to the open form. The explanation given for the lack of further change was that binding of the first copper caused the pre-organization of the second binding site and therefore no further conformational change was required.

A binuclear ascidiacyclamide Cu^{2+} complex has been crystallized and shows the presence of a μ-carbonato bridge between the metal centers (Figure 25).[69] This crystal structure shows explicitly the positions occupied by the copper atoms and identifies them unambiguously as being 5-coordinate with distorted square pyramidal/trigonal bipyramidal (SQP/TBP) geometries, the peptide part of the primary coordination sphere consisting of three nitrogen atoms. The nitrogens involved are donated by the thiazole, N_T, the (isoleucine) amide, N_A, and the oxazoline, N_O, on either side of the macrocycle. For this reason the patellamide binding site has become known as the TAO motif.[20] The other positions are occupied by oxygens, one from the carbonate bridge and the other from a coordinated solvent molecule. This general motif is assumed for the other patellamides and this has been supported by EPR [70] and MS [68] analysis of the di-copper complexes of patellamides A, C and D. No more crystal structures have been published for the natural patellamides but some of the synthetic analogues mentioned above have been obtained in crystalline form.[16,72] These have not suggested anything that might contradict the view that the structure for any of the other patellamides is likely to be identical to that shown for ascidiacyclamide.

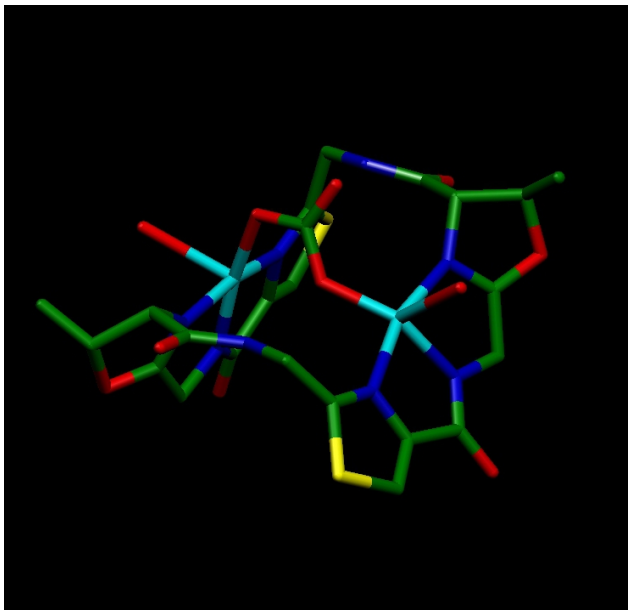


Figure 25. X-Ray structure of ascidiacyclamide dicopper(II) complex.[69] Hydrogen atoms and valine and isoleucine side-chains have been removed in order to improve clarity

Co^{2+} and Ni^{2+}

Both of these metals were studied in conjunction with patellamides A and C.[68] Neither peptide displayed any detectable affinity for either Co^{2+} or Ni^{2+}. This result was unexpected as both metals are as common in biological systems as Cu^{2+} and Zn^{2+}. Because of this it had been thought that the patellamides should show some complexing ability for these metal ions, even if it was non-specific in nature. These experiments did, however, highlight the selectivity exhibited by the patellamides for different metal ions from the first transition series.

It was thought that the inability of either Co^{2+} or Ni^{2+} to effect deprotonation of the amide nitrogen in the TAO binding site might be responsible (at least in part) for this result.[68] In addition, preference for octahedral coordination geometry might have made complexation less likely due to steric hindrance within the binding site caused by the presence of the amino acid side-chains.[68] The importance of the second point is, perhaps, debatable as the coordination numbers and geometries of the two metals in their divalent states are known to be highly variable and display a strong dependance on the nature of the ligands involved.[75,77]

A possibility not addressed by the authors is that the solvation properties of Co^{2+} and Ni^{2+} may be considerably different to those of Cu^{2+} and Zn^{2+}. This would result in different relative stabilities of the metal-solvent and metal-patellamide complexes (and consequently different equilibria between these species in solution). This is supported by the logK values for solvent exchange in water of the four metal ions which have been reported as ~6.3, ~4.5, ~9.6 and 7.5 for Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+} respectively.[79] Values for exchange of methanol (which was the experimental solvent in the complexation studies) at Co^{2+}, Ni^{2+} and Cu^{2+} are also given by this source as ~3.9, ~2.6 and ~8 respectively.

Reactivity of patellamide dicopper complexes

It has been suggested by several authors that the presence of ligands such as the bridging carbonate in the ascidiacyclamide dicopper complex[69] (Figure 25) may in fact be the result of catalytic activity in the metal center itself.[68,70,71] Van den Brenk *et al.*, 1994,[70] found that when the dicopper complex of patellamide D was incubated in methanol with 2 equivalents of base (triethylamine (Et_3N)) the largest peak (m/z=945.0) in the mass spectra of the resulting solution corresponded to the carbonate adduct of the complex. Increasing the incubation period or the concentration of base lead to a corresponding increase in the magnitude of the peak. As no carbonate had been added to the solution, the authors suggested that CO_3^{2-} was being formed either by direct action of the dicopper center on absorbed atmospheric CO_2 or alternatively by nucleophilic attack on the CO_2 by hydroxide ion produced as a result of Et_3N induced deprotonation of water present in the methanolic solution. Further investigations involving titrations monitored by electronic absorbtion, CD and EPR spectrometry supported the former mechanism.[70] Although not pointed out by the authors, the common occurance of hydroxide adducts of the dicopper complex (detected in their mass spectral analyses) also seems to fit in with the copper-based mechanism.

In studies performed on patellamide C by Morris and co-workers[68,71] a similar phenomenon was observed with the MS of the methanolic solution of patellamide C and $CuCl_2$ displaying a peak at m/z=975. This was attributed to an adduct with the ligand formula $C_3H_6O_3$. It was suggested that the ligand species might be dimethylcarbonate produced by the reaction of CO_3^{2-} (formed from atmospheric CO_2) with methanol. In these

studies no base had been added to the solutions (as was the case in the work on patellamide D)[70] and this was taken as further support for the possibility that the dicopper complex was acting as a catalyst for the reaction. The rate of formation of this adduct was found to be much faster when copper acetate was used in place of the chloride.

Potential biological significance of patellamide/copper interaction

The apparent catalytic behaviour of the dicopper complexes of ascidiacyclamide and patellamides C and D was taken to suggest a possible biological role for these species.[68,70,71] This is supported by the fact that the intercopper distance in the patellamide core (3.6 – 4.5 Å)[70] is similar to that observed in oxygen activating dicopper enzymes such as phenoloxidase and tyrosinase, both of which catalyse the hydroxylation and subsequent dehydrogenation of phenolic compounds.[77,80,81] Synthetic dicopper complexes have also been shown to form dimethylcarbonate in methanolic solution with the maximum rate of formation occuring when the intercopper distance is approximately 4 Å.[82]

In their speculative discussions of the possible biological significance of patellamide/copper complexes, the authors cited above have neglected to mention whether there is any evidence of the involvement of copper in ascidian physiological processes. Inspection of the literature during the course of the present work revealed that research into the effects of this metal on important aspects of ascidian biology has been ongoing since the early part of the twentieth century. The results of these studies are summarised in the following sections.

Copper in the developmental stages of the ascidian lifecycle

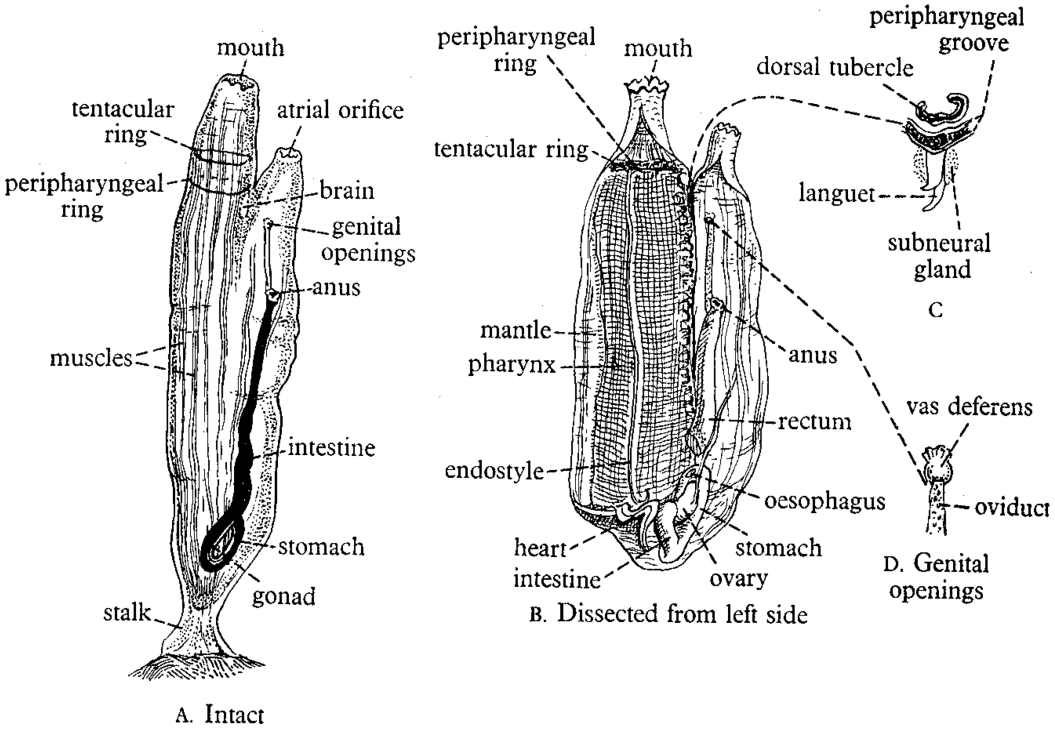


Figure 26. The solitary ascidian *Ciona intestinalis*. A, Intact organism. B, Dissection showing internal distribution. C, Detail of the environment of the brain. D, Detail of the genital openings within the atrium. Reproduced from Borradaile (1963), p722.[2]

Ascidians are hermaphroditic organisms with both male and female reproductive organs found adjacent to the digestive tract. In solitary species such as *Ciona intestinalis* (order *Enterogona*, family *Cionidae*, Figure 26) fertilisation occurs externally in the surrounding water after the release of eggs and sperm via the atrial orifice whereas in colonial species (such as *Lissolinum patella,* order *Enterogona*, family *Didemnidae*) fertilisation and incubation of the egg occurs inside the individual ascidian.[1] The fertilised egg (Figure 27, **1**) develops into a free-swimming larva or tadpole (Figure 27, **2**) and it is at this stage that the organism takes on the morphology that most strongly places it in the phylum urochordata (pseudo-vertebrates). The tadpole has in its tail section a column of vacuolated cells known as the notochord that contains or is associated with a nerve-cord; these features are morphologically and functionally similar to both the spinal column and spinal cord in the "true" vertebrates (Figure 27, **4**).[1]

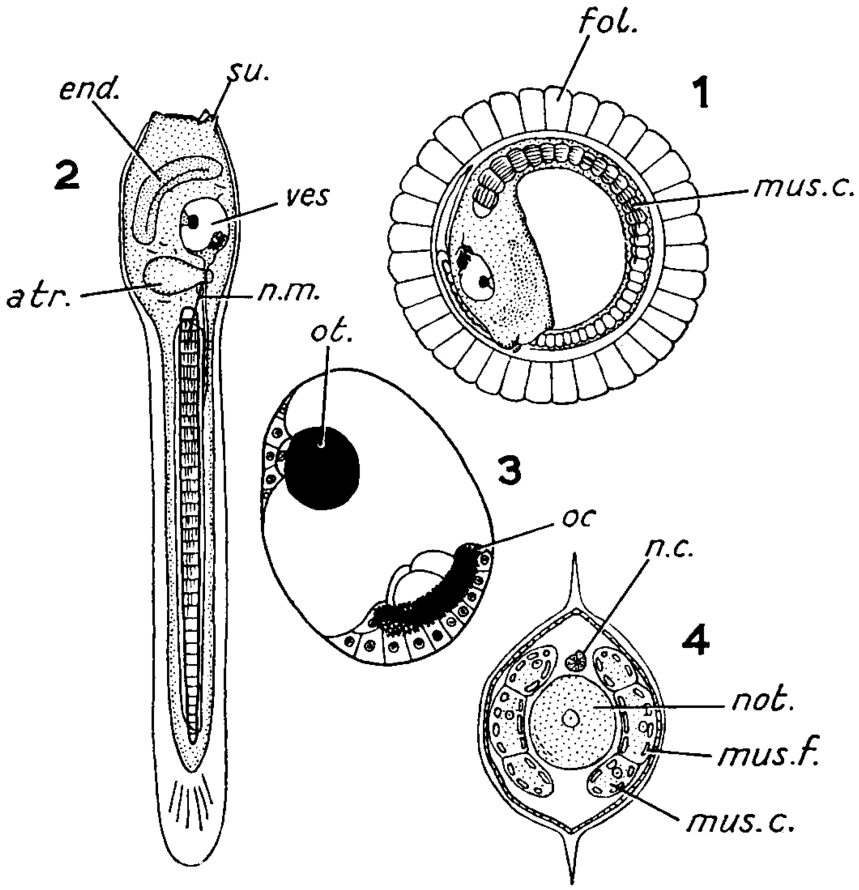


Figure 27. Development of the ascidian tadpole. 1, Tadpole ready to hatch. 2, Tadpole. 3, Sensory vesicle. 4, Cross section of the tail. Reproduced from Young (1962), p68.[1] Abbreviations: *atr.* atrium; *end.* endostyle; *fol.* follicle cells; *mus.c.* muscle cells; *mus.f.* muscle fibrils; *n.c.* nerve-cord; *n.m.* nerve to tail muscles; *not.* notochord; *oc.* ocellus; *ot.* otolith; *su.* sticking gland; *ves.* sensory vesicle.

The larvae contain both the temporary organs required for this phase of development and the rudiments of the permanent structural features needed later in the adult animal. This free-swimming period varies in duration between species but comes to an end identically when the larva attaches itself to a suitable surface and enters the phase of its development known as metamorphosis (Figure 28). This metamorphosis sees the tadpole changing into its adult form and begins with the loss of its tail section and the almost complete decomposition of its internal structures into a mixture of materials from which the previously dormant rudiments can begin to assemble the components necessary for the development of the adult morphology.

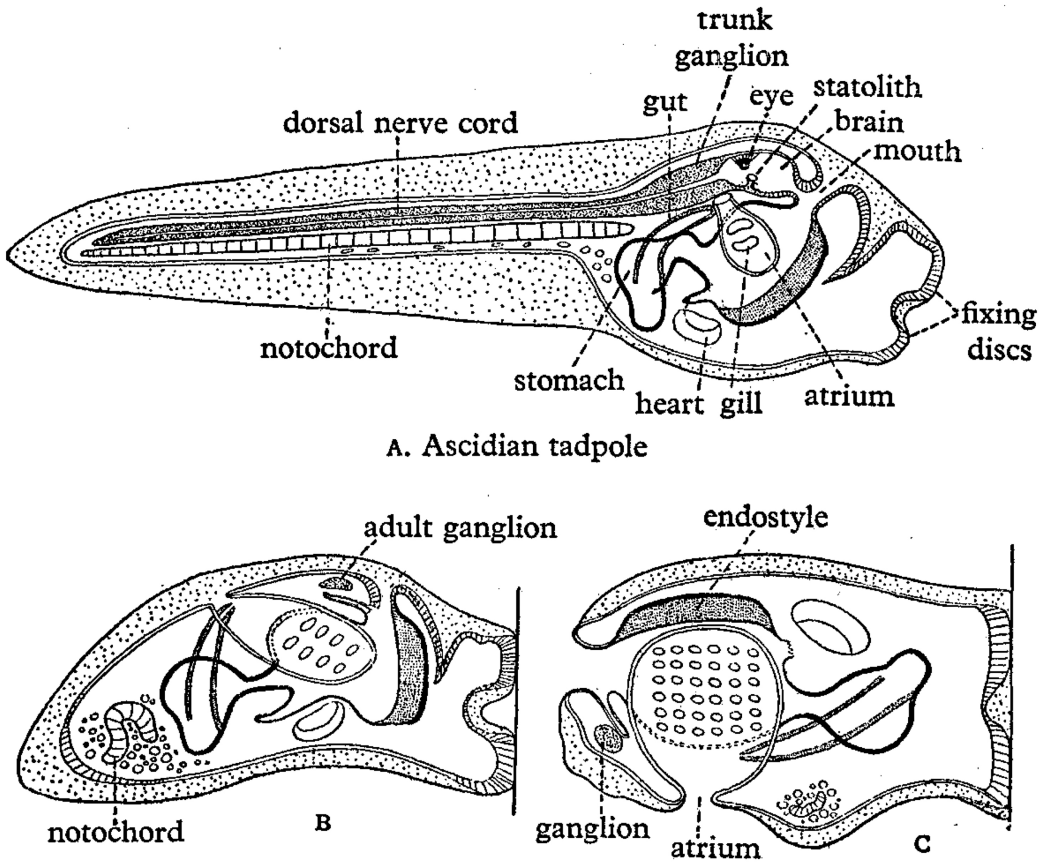


Figure 28. Diagram of the metamorphosis of an ascidian tadpole larva. A, At time of fixation. B, Midway in metamorphosis. C, Metamorphosis complete. Reproduced from Borradaile (1963), p725. [2]

Investigations into this process in the first half of the 20th century by Grave and co-workers (references 8-19 in Glaser and Anslow, 1949)[83] highlighted the fact that the development of ascidian offspring was susceptible to external environmental influences. Adult organisms, eggs, early embryos and tadpoles of the species *Ascidia* (solitary ascidian; order *Enterogona*, family *Ascidiidae*) and *Polyandrocarpa* (solitary ascidian; order *Pleurogona*, family *Styelidae*) were pulped in sea water with Berkshire sand as an abrasive and the filtrates of the resulting mixtures used to probe for inhibitory or stimulatory effects in cultures of eggs and larvae. It was found that the filtrates from the adults and tadpoles, as well as some of the more mature embryos, acted as strong promoters of the onset of the metamorphic stage of development. Dissolved inorganic materials came to be suspected as being possible mediating factors due to the observation that some of the controls in their experiments, which contained only the sand and sea water, showed effects similar to the filtrates.

Further investigations were carried out using solutions of heavy metals in place of the pulped organism filtrates. Only solutions of $FeCl_2$ and $CuCl_2$ showed effects similar to those obtained with the filtrates. The experiments showed that for *Polyandrocarpa* the time before the disruptive metamorphic stage was reached was 8 minutes with exposure to $CuCl_2$ at $3*10^{-6}$ M but only 2 minutes when the sea water with 0.1% chloretone (chlorobutanol) and $4*10^{-6}$ M $CuCl_2$. The maximum and minimum swimming period prior to attachment for *Styela* larvae were reduced from the usual 100 hours and 9 hours to 2.5 hours and 30 minutes, respectively.

Glaser and Anslow continued this line of investigation and in their 1949 paper[83] suggested that a mechanism existed in which Cu-containing redox enzymes involved in the maintenance of the body of the larva were in competitive equilibrium with secondary Cu-binding chemical species (either secondary sites on the enzymes or their substrates). Inactivation of these enzymes was postulated as leading to the destructive phase in which the body of the larva breaks down in preparation for the activation of the rudiments. The constructive phase occuring after this was attributed to activation of enzymes in the rudiments that required Cu for proper functioning. Two possible routes by which the activation of the rudiment enzymes could occur were proposed. The first of these involved competetive absorption by the larval systems followed by activation once these systems had become saturated. The second involved the saturation of both primary and secondary sites in the rudiments early in the development of the larva causing these systems to remain inactive until at a later time the Cu concentration began to fall to levels conducive to their reactivation.

More recently, Bellas *et al*., 2001, investigated the effects of differing concentrations of Cr, Cu, Cd and Hg on the early stages of the reproductive cycle of *C. intestinalis* in a study related to the development of water quality assays.[84] Stock solutions of the metals (analytical grade $Hg(II)Cl_2$, $Cu(II)Cl_2.2H_2O$, $Cd(II)Cl_2.H_2O$ and $Cr(VI)O_3$) were prepared in double-distilled water before being added to suspensions of biological material in artificial seawater. Glassware was acid washed (HNO_3, 10% vol.) and rinsed with double-distilled water prior to use. It was found that in assays involving embryonic development and the rate of attachment prior to metamorphosis, Hg was 2.7 times more toxic than Cu,

approximately 30 times more toxic than Cd and nearly 1000 times more toxic than Cr. The effective concentrations that reduced the frequency of both events by 50% (EC_{50}) were [Hg]=$2.7*10^{-7}$ M, [Cu]=$7.2*10^{-7}$ M, [Cd]=$7.5*10^{-6}$ M and [Cr]=$2*10^{-4}$ M.

Copper and the ascidian immune response

Toxicity of copper towards cellular processes in adult ascidians was found in a study performed on *Styela plicata* (solitary ascidian; order *Pleurogona*, family *Styelidae*) by Raftos and Hutchinson, 1997.[85] In this work, intact organisms exposed to copper concentrations of $7.9*10^{-2}$ M or more displayed acute toxicity indicated by failure of the organism to react to physical stimulation of the atrial opening and the results of haemocyte analyses. All animals exposed to this level of copper died within eight days of the beginning of the experiment. Exposure to copper at lower concentration ([Cu]=$7.9*10^{-6}$ M) resulted in reduction of haemocyte proliferation by 82% and reduction of the phagocytic activity of these cells by 62%.

Similar studies have highlighted the ability of copper to effect changes in phenoloxidase (PO) activity in *S. plicata*,[85-88] *Halocynthia roretzi*[89] (solitary ascidian; order *Pleurogona*, family *Pyuridae*) and *Botrylloides leachi*[90] (colonial ascidian; order *Pleurogona*, family *Styelidae*). PO is the final product of an enzyme cacade involved in the immune response of ascidians and is localised in, and released from, specalised vacuolar haemocytes known as morula cells.[86,88] PO catalyses the orthohydroxylation of monophenols such as tyrosine and their subsequent conversion to orthoquinones. These quinones then polymerise without further catalysis to form melanin.[86] The metallo-enzyme is itself produced by activation of the apo-protein prophenoloxidase (proPO) by the binding of two copper atoms in active sites similar to those found in the arthropod dioxygen transporting enzyme haemocyanin (HC).[81]

The copper dependance of PO function was demonstrated by Hata *et al.*, 1998.[89] In this study the activity of purified PO from haemocytes of *H. roretzi* was shown to be strongly inhibited by "various concentrations"[89] of phenylthiourea or diethyldithiocarbamate. Cysteine and dithiothreitol (*threo*-2,3-dihydroxy-1,4-dithiolbutane) showed a lesser degree of inhibition. The inhibition by phenylthiourea and diethyldithiocarbamate was shown to

be completely reversed by the addition of Cu^{2+} but the metals Mg^{2+}, Ni^{2+} and Zn^{2+} had little effect. The common chelating agent EDTA was found to have little effect on the activity of the purified enzyme but was able to suppress the expression of PO in whole haemocytes exposed to a range of stimuli known to cause its release. These included whole sheep erythrocytes, yeast cells, zymosan (crude preparation of yeast cell walls containing mainly protein-carbohydrate complexes) and lipopolysaccharide.

The fact that PO activity was increased on exposure to these agents whilst no copper was added to the stimulation experiments suggests that the necessary Cu^{2+} might have been released from stores within the cell when required. This is reminiscent of the suggestion by Glaser and Anslow that an equilibrium between Cu-containing redox enzymes and some secondary binding agent was involved in the regulation of metamorphic changes in ascidian tadpoles (previous section).[83]

Chapter 4

Density Functional Study of Cu^{2+} Binding in the Patellamide TAO Binding Site

Previous computational studies on patellamide Cu^{2+} complexes

The application of computational methods to the study of patellamide/metal complexes has been very limited in comparison with the experimental studies discussed above. Several authors have made use of nOe- [20,71,91] and EPR- [73,91] restrained dynamics to study the effects of metal complexation on the global conformational properties of the patellamides. This approach has been very useful in visualising data obtained in the experimental studies above but the fact that the location of the metal centers and their effects on the peptide must be inferred means that they do not contribute a great deal of information about the complexation itself. It is interesting to note that the two methods in fact provide complimentary information, with details of the peptide geometry and the copper-copper distance available from the nOe and EPR data respectively. Combining these data might give a more complete description of complexes of this type.

The only study found in the literature that explicitly included the complexed metal was made by Comba *et al*. who used their own force field (MOMEC)[92] to study the effects of copper binding on a synthetic patellamide derivative.[91] The peptide, called PatN (**26**) by the authors, was similar to patellamide A in that it retained the isoleucine residues and the serine and threonine residues which form the oxazoline rings in the natural compound. Unlike patellamide A, however, the oxazoline rings remained uncyclised and the valine residues adjacent to the thiazole rings were replaced by glycine (Figure 29). The authors stated that these changes were made in order to avoid the problem of epimerization of asymmetric centers.

In the force field model the copper centers were represented as six-coordinate octahedra with three equatorial sites filled by the same backbone nitrogens as were seen to be complexed in the crystal structure of ascidiacyclamide shown above. The remaining equatorial site and the two axial sites were occupied by aquo ligands. This model was used to in conjunction with EPR data obtained as part of the same study and indicated a solution structure broadly similar to that of the crystal phase ascidiacyclamide/Cu_2CO_3 complex (Figure 25). The main difference was that the structure was considerably flattened compared with the distinct saddle-like shape of the crystal structure. This effect may, however, have been an artefact of the force field parameterisation as in the

ascidiacyclamide structure the copper was five-coordinate with the second axial position hindered by the isoleucine side-chains and apparently inaccessible to solvent.

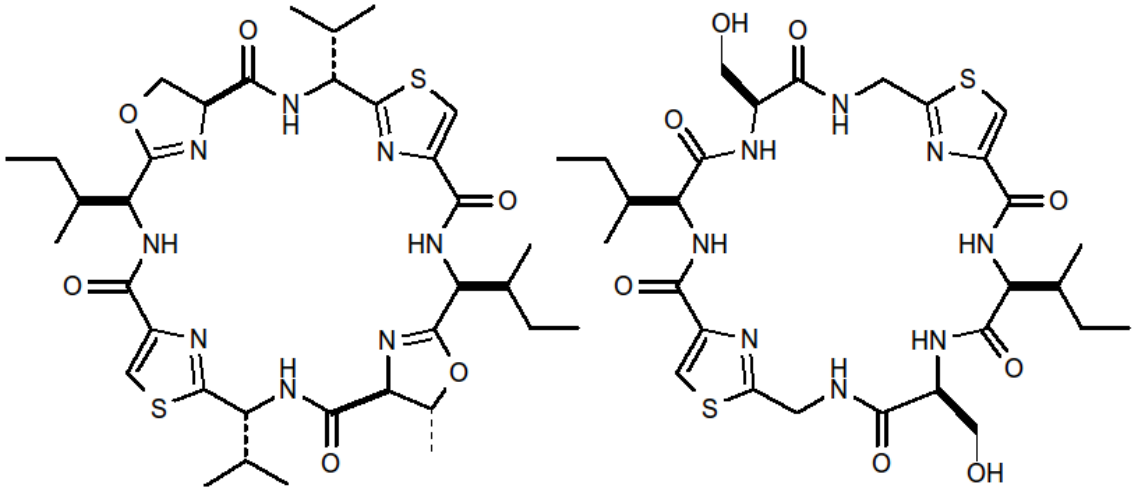


Figure 29. Structures of natural patellamide A and the synthetic derivative PatN.

A further complication in attempts to rationalise the deviation from the ascidiacyclamide crystal structure arises from the substitutions made in the synthetic compound which caused the removal of two of the characteristic features of the patellamides, namely the corner-forming oxazoline rings and the two D-amino acid residues. The lack of the valine residues removes the possibility of steric interactions across the macrocycle and in combination with the failure to form the oxazoline rings introduces several degrees of freedom into the macrocyclic structure that are not available to the natural compounds.

In light of the demonstrated importance of copper in ascidian biochemistry and the experimental findings relating to the copper complexes of the patellamides, it was decided to investigate the nature of these complexes using quantum chemical methods. This approach does not suffer from the problems inherent in the molecular mechanics techniques used previously and consequently it was hoped that new information might be obtained that would add to the existing work and perhaps suggest some possibilities for future investigations.

The reactivities and other properties of transition metal complexes are known to depend to a large extent on the nature of the ligands attached to the metal and the geometry of the resulting complex.[74] In order to investigate both of these influences, geometry optimisations and subsequent calculation of equilibrium properties relating to the reactivity of the copper center were performed with a number of small ligand species observed in the experimental works.

Advantages/disadvantages of different electronic structure methods in the study of transition metal complexes

The application of conventional quantum chemical methods to the study of metals from the first transition series has long been recognised as being highly problematic. This stems from the large number of electrons that must be treated and from the fact that electron correlation in these large systems becomes even more important than in 1st- and 2nd-row main group molecules due to the presence of filled d-orbitals. In the few cases where high accuracy experimental data exists for TM complexes (and these are generally closed-shell) it is found that the Hartree-Fock (HF) method systematically overestimates M-L bond lengths whilst the simplest correlated post-HF method, MP2 (Møller-Plesset second order perturbation theory), gives a systematic underestimation.[39,93] Interestingly, the poor performance of the MP2 method seems restricted to the 1st transition series with calculations on systems containing metals from the 2nd and 3rd transition series providing considerably more accurate results.

More computationally demanding post-HF methods such as CCSD(T) (coupled-cluster single- and double-excitation method with noniterative inclusion of triple excitations) have been shown to give much improved M-L bond lengths but the increase in CPU time resulting from the use of such methods means that on computer systems with low to moderate capabilities calculations become unfeasibly large, even for combinations of metals from the left-hand side of the 1st transition series with small ligands such as CO.[93] As the metal/ligand combinations that are of interest in a biological context are usually much larger than this and may contain atoms from later in the main group such as phosphorus and sulfur, theoretical studies on biologically relevant complexes at this level are limited and those that do exist are often single-point calculations performed at some

fixed geometry (possibly optimized at a lower level of theory or obtained from experimental results).

Density functional theory (DFT) has come to the fore in recent years as a valuable alternative to conventional correlated methods in studies of transition metal-containing systems. The use of hybrid density functionals containing an admixture of HF exchange has raised the accuracy that may be routinely obtained with the DFT approach to a level comparable with methods such as CCSD(T) but at a fraction of the computational cost. Furthermore, DFT has the added attractions that it is size-extensive, and may therefore be applied equally rigorously to both large and small systems, and that it may be applied to any area of the periodic table with equal confidence.

A final, but operationally very important, point relating to the choice of DFT is the performance of the model in open-shell situations. The HF model (and consequently any post-HF corrections added to it) suffers from the major drawback that calculation of an unrestricted HF wave function is complicated by the occurrence of a phenomenon known as spin contamination. This is due to mixing of the wave functions of the state of interest with those corresponding to other states of differing multiplicity to produce unphysical hybrid states (because the number of electrons, N, is conserved only configurations obtainable through rearrangement of the N electrons will be involved e.g. a doublet wave function might suffer from contamination by the higher-lying quartet). Perhaps the most immediately important adverse effect of such contamination is that spin-contaminated wave functions are generally much more difficult to converge and it is common to find that in UHF calculations the energy will oscillate rather than smoothly approaching a stable value and the calculation must often be abandoned for this reason alone.

The unrestricted Kohn-Sham (UKS) formalism in DFT is considerably less susceptible to spin contamination and is therefore much easier to use than the UHF method. Furthermore, the greatly reduced contamination occurring for the DFT variant means that the results obtained with this method are more reliable in that they can be taken to relate to the state of interest rather than to an unidentifiable mixture of states. This is a very desirable property when considering transition metal containing systems because of the

common occurrence of open-shell oxidation states for these metals.

Selection of a model of the Cu^{2+} binding site

Despite the favourable scaling properties of density functional calculations, modelling the patellamide/Cu^{2+} interaction using the entire peptide structure was not feasible at the chosen level of theory with the available computational facilities. As Cu^{2+} was known from the previous experimental work to specifically bind to the nitrogens of the TAO motif,[68-70] a model of the binding site could be employed to represent the coordination environment in the whole peptide which only consisted of those portions of the molecule likely to have a significant effect on the electronic environment of the ligand atoms. The model binding site selected (Figure 30) comprises what is effectively half of the patellamide molecule.

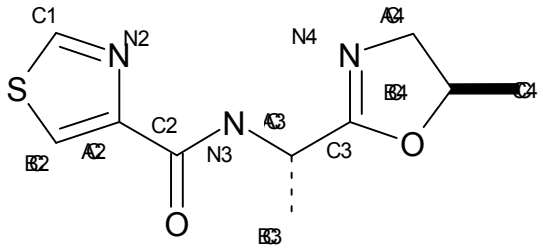


Figure 30. Model structure used to represent the TAO binding site found in the patellamides.

The amino acid residue side chain attached to the alpha carbon of the central amide was replaced by a simple methyl group so that the model represented a generalised patellamide rather than a specific species whilst retaining the possibly important structural contribution from the substitution at this point in the site. It was decided that any changes made to the thiazole and oxazoline rings should be as minimal as possible. The extensive electronic delocalisation involving the thiazole ring and the adjacent amide unit meant that both were likely to be fully involved in the binding process and modification of the oxazoline sub-unit such as ring-opening had been shown experimentally to have a major effect on the nature and properties of the peptide.[72] Furthermore, the methyl substituent on the oxazoline ring might also have some effect both on the conformation of the ring as well as possibly interacting with any exogenous ligands attached axially to the copper. It was necessary also to decide where to truncate the model with regard to the "bridges"

connecting the two binding sites (that is, the portions containing the alanine and phenylalanine residues in e.g. Patellamide D or the two valine residues in e.g. ascidiacyclamide). The simplest way to do this was to replace the macrocyclic carbons attached to both the thiazole and oxazoline rings with hydrogens. The model site selected was therefore deemed to be the smallest that it was possible to use whilst still retaining what appeared to be the key features and so remaining physically justifiable.

The ability of the model site to reproduce the geometry of the equivalent portion of the complete molecule was assessed by optimising the fragment at the B3LYP/SBKJC(d) level (chosen to be used in the metal-containing systems – see next section). The calculation was performed using GAMESS-UK v6.3 [30] on a dual 800MHz pentium workstation with 1Gb RAM running Red Hat linux (v7.0) as were all subsequent calculations on the copper-containing TAO systems.

The results of this calculation were compared with the average geometry (bond lengths and angles) obtained from the available crystal structures for natural patellamides in the open conformation (Table 19).[13,14,26-28,94] Despite the lack of the remainder of the molecule, the geometry calculated at this level of theory gave bond lengths differing from the average by 0.6 - 2.5% (0.5 – 2.9 standard deviations (SD)). Bond angles were reproduced to within 0.1 – 1.5% (0.0 – 1.8 SD). Perhaps most importantly, the geometry obtained for the TAO site itself (distances between thiazole (N_T), amide (N_A) and oxazoline (N_O) donor nitrogens) was N_T-N_A = 2.806 Å (0.7% / 0.3 SD), N_A-N_O = 2.768 Å (0.2% / 0.1 SD) and N_T-N_O = 4.559 Å (1.7% / 0.5 SD). The angles calculated for the TAO site were N_T-N_A-N_O = 110.0º (2.7% / 0.8 SD), N_A-N_T-N_O = 34.8º (4.3% / 0.7 SD) and N_A-N_O-N_T = 33.5º (5.0% / 0.9 SD) The largest deviations in bonds and angles were seen to occur at the thiazole sulfur, the backbone bonds and angle at the carbonyl carbon and in the carbon-carbon single bond of the oxazoline ring. Overestimation of the sulfur-carbon bond lengths was expected for this combination of basis set and density functional (see results of thiazole optimisations – Chapter 1). The variation in the single carbon-carbon bond of the oxazoline unit was not unexpected due to the fact that the equivalent bond in the full patellamides was substituted at one end by the amide carbonyl of the bridging portion of the macrocycle rather than a hydrogen atom as in the model. Whilst the deviation around

the amido carbon occurred in a position that would be directly adjacent to a donor atom in the copper complex it was decided that in light of the good accuracy obtained for the TAO site itself that this would not be likely to cause any major errors in the metal-complexed structure.

Bond	Mean (exp.)	SD (exp.)	Model	Error	% Error	# of SDs	Angle	Mean (exp.)	SD (exp.)	Model	Error	% Error	# of SDs
1C-2N	1.300	0.020	1.311	0.011	0.9	0.6	1C-2N-2CA	111.4	1.3	110.7	-0.7	-0.6	0.5
2N-2CA	1.379	0.013	1.391	0.012	0.9	1.0	2N-2CA-2C	121.4	1.6	121.6	0.2	0.2	0.1
2CA-2C	1.479	0.017	1.516	0.037	2.5	2.2	2CA-2C-3N	115.9	1.3	114.2	-1.7	-1.5	1.3
2C-3N	1.339	0.015	1.369	0.030	2.2	2.0	2C-3N-3CA	120.4	1.8	121.4	1.0	0.8	0.6
3N-3CA	1.449	0.020	1.476	0.027	1.9	1.4	3N-3CA-3C	108.4	1.6	108.5	0.1	0.1	0.0
3CA-3C	1.498	0.009	1.524	0.026	1.7	2.9	3CA-3C-4N	126.6	0.9	125.9	-0.7	-0.6	0.8
3C-4N	1.251	0.026	1.286	0.035	2.8	1.4	3C-4N-4CA	106.8	0.7	106.3	-0.5	-0.5	0.7
4N-4CA	1.472	0.011	1.481	0.010	0.6	0.9							
1C-S	1.736	0.019	1.759	0.023	1.3	1.2	1C-S-2CB	90.6	0.8	89.1	-1.5	-1.6	1.8
S-2CB	1.696	0.020	1.736	0.040	2.4	2.0	S-2CB-2CA	109.4	0.9	109.7	0.3	0.2	0.3
2CB-2CA	1.357	0.019	1.391	0.034	2.5	1.8	2CB-2CA-2N	115.6	0.7	115.5	-0.1	-0.1	0.2
2C=O	1.233	0.018	1.244	0.011	0.9	0.6	2CA-2C-O	120.7	1.0	121.7	1.0	0.8	1.0
4CA-4CB	1.532	0.014	1.568	0.036	2.3	2.5	3C-O-4CB	105.7	1.6	106.1	0.4	0.3	0.2
4CB-O	1.451	0.023	1.476	0.025	1.7	1.1	O-4CB-4CA	103.2	1.6	102.9	-0.3	-0.3	0.2
O-3C	1.364	0.018	1.374	0.010	0.7	0.5	4CB-4CA-4N	104.6	1.1	105.5	0.9	0.8	0.8
N_T-N_A	2.787	0.058	2.806	0.019	0.7	0.3	N_T-N_A-N_O	113.1	3.9	110.0	-3.2	-2.7	0.8
N_A-N_O	2.772	0.048	2.768	-0.004	-0.2	0.1	N_A-N_T-N_O	33.4	2.2	34.8	1.4	4.3	0.7
N_T-N_O	4.637	0.172	4.559	-0.078	-1.7	0.5	N_A-N_O-N_T	33.5	1.8	35.2	-1.7	-5.0	0.9

Table 19. Comparison of mean geometrical parameters obtained from patellamide crystal structures with geometry of model TAO site (B3LYP/SBKJC(d)). Bond lengths in Å and angles in degrees.

Preferred coordination geometries

Optimisation of the model TAO-Cu^{2+} complex was performed with a number of exogenous ligands thought to be of importance in both the natural and experimental environments of the patellamides. The B3LYP hybrid density functional was used throughout.[41]

This functional has been used a number of times in studies on copper complexes and has been shown to give a high degree of accuracy when compared both with experimental geometries and other DFT and wavefunction based methods.[95-103] It is important to note,

however, that in the studies that included comparison with a number of other density functional types B3LYP did not in fact display the best overall performance.[97,102] Despite the slight improvement over B3LYP shown by functionals such as SVWN (local density approximation (LDA)) and BLYP (generalised gradient approximation (GGA)), the performance of the hybrid functional was still considerably better than than that of either the HF or MP2 methods (using double-ξ quality basis sets).[97,102]

In the studies dealing with large ligands (of similar size to the model TAO ligand or larger) the maximum average deviations in bond lengths and angles involving copper when using B3LYP were -0.089 Å and -2.2º [101] This was obtained with the rather small 3-21G(d,p) basis but when sets of similar quality to 6-31G(d,p) were used the average deviations dropped to within ~0.02 Å and ~1.3º of the experimental values.[97,98,100,101,103] In light of this information and the performance of B3LYP in the thiazole/oxazole and full patellamide studies in the present work, it was decided that the continued use of this functional was justified.

The relativistic effective core potential basis set of Stevens *et al.* was employed for these calculations.[48] This basis uses a -31G splitting of the valence space (referred to here as SBKJC-31G) for main group atoms and a -311G split for transition metals (SBKJC-311G). The quality of this basis in the valence region is similar to that of the commonly used all-electron 6-31G set but has the added advantage that the core orbitals are removed and replaced with the effective potential therefore reducing the total number of functions required in the calculations. This was most important for the Cu and S atoms where the number of electrons treated explicitly was reduced from 27 and 16 to 17 and 6 respectively. In the other main group atoms only 2 electrons (the 1s pair) were removed but the size of the TAO ligand meant that a total of 30 electrons were removed in addition to the 20 from Cu and S. Polarisation functions were expected to be important in the TAO ligand because of the extensive π-density and for the main group atoms a d-function was added to each yielding the SBKJC-31G(d) basis. Additional hydrogen p-functions were neglected as these were unlikely to contribute much to the heavy atom geometry (as was found in the study of basis set effects in thiazole and oxazole - Chapter 1). The cationic Cu^{2+} was left unpolarised with the SBKJC-311G basis.

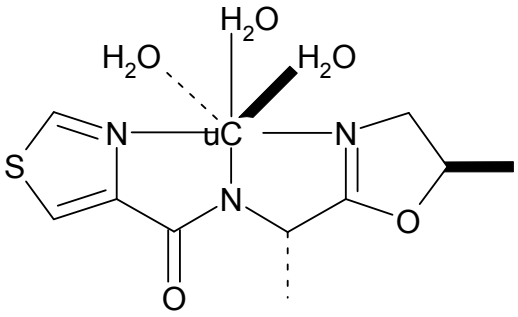


Figure 31. Starting configuration for $[Cu(TAO)(H_2O)_n]^+$ optimisation.

Neutral ligands

The coordination of aquo ligands to the copper center of the TAO-Cu^{2+} complex was investigated as these represent the most abundant hetero-ligands that might be expected to become involved with a biological system such as the patellamide-Cu^{2+} complex. As the experimental studies on the copper complexes of the patellamides had been performed in methanol[68-71] this ligand was included in the present study for the purpose of comparison with the aquo ligand case.

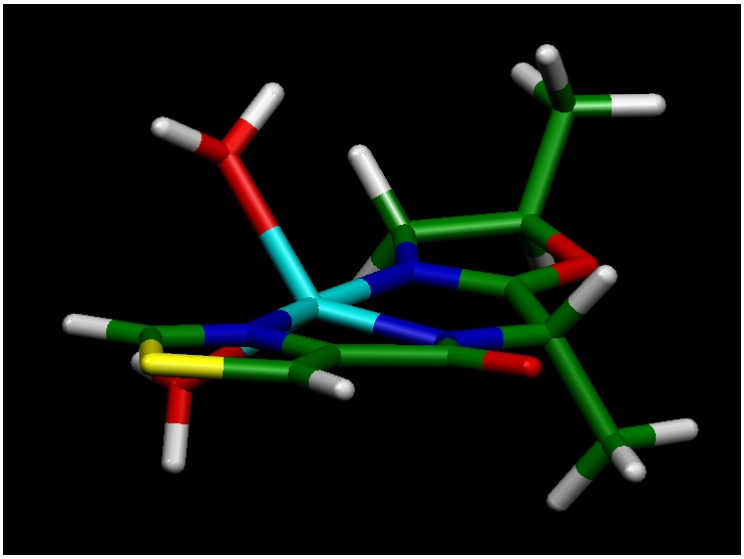


Figure 32. $[Cu(TAO)(H_2O)_2]^+$ equilibrium geometry calculated at B3LYP/SBKJC level (d-functions on main group atoms).

$[Cu(TAO)(H_2O)_3]^+$

The model used as input in this calculation consisted of the TAO-Cu^{2+} unit complexed to three aquo ligands (Figure 31). Optimisation of this structure led to the dissociation of the aquo ligand extending down into the plane of the paper (in Figure 31), resulting in a distorted trigonal bipyramidal coordination geometry with the dissociated water molecule

hydrogen bonded to the (previously) equatorial aquo ligand. The dissociated water was removed and the structure re-optimized but only a small rearrangement was observed. The final coordinatively unsaturated geometry is shown in Figure 32 and the values obtained for the bond lengths and angles are given in Tables 20 & 21.

$[Cu(TAO)(MeOH)_3]^+$

The optimisation performed for the $[Cu(TAO)(H_2O)_3]^+$ complex was repeated with three methanol molecules as the hetero-ligands in order to establish whether this solvent had any significant effect on the coordination geometry of the copper center relative to the aquo ligand case. A similar 5-coordinate geometry resulting from dissociation of the lower axial ligand (Figure 33) was observed with the trigonal bipyramidal distortion of the square pyramid essentially unchanged from that of $[Cu(TAO)(H_2O)_2]^+$ (see Figure 32).

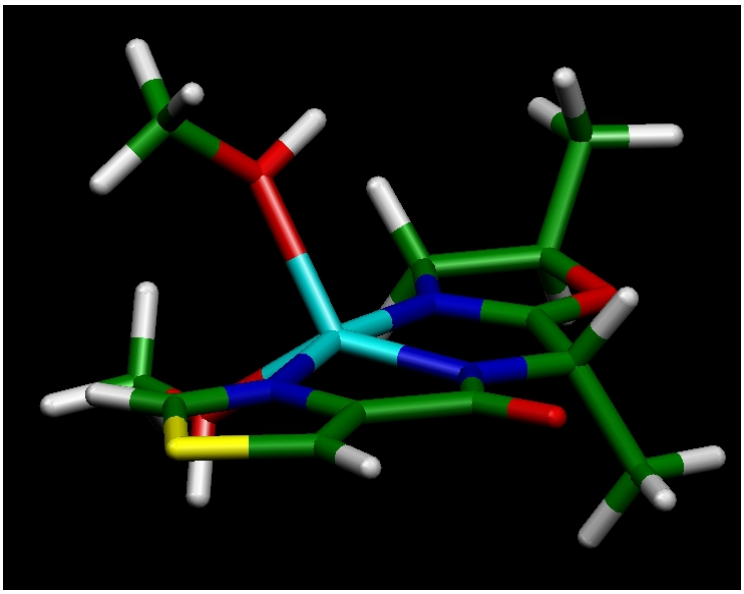
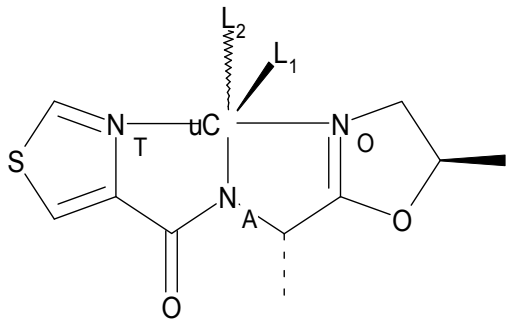


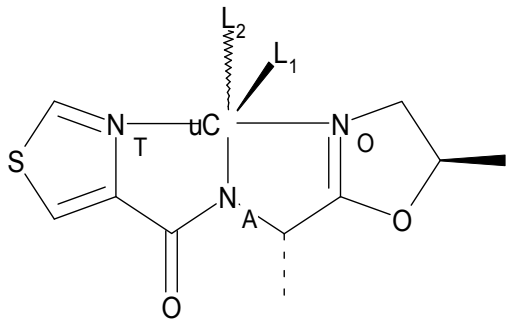
Figure 33. $[Cu(TAO)(MeOH)_2]^+$ equilibrium geometry calculated at B3LYP/SBKJC level (d-functions on main group atoms).

Again, removal of the dissociated ligand and subsequent re-optimisation caused only minor changes in the final geometry. The deviations in bond lengths in the primary coordination sphere with respect to those in the $[Cu(TAO)(H_2O)_2]^+$ complex were of the order of 0.02 Å (Table 20) Bond angles deviated from the aquo complex by no more than 0.5° (N-Cu-N angles) and by 1.3° and 5.2° for the angles involving the lower and upper (see Figure 33) MeOH ligands, respectively (Table 21).



Bond	Mean (exp.)	SD (exp.)	$[Cu(TAO)(H_2O)_2]^+$	% error	# of SDs	$[Cu(TAO)(MeOH)_2]^+$	% error	# of SDs
1C-2N	1.333	0.033	1.326	-0.5	0.2	1.326	-0.5	0.2
2N-2CA	1.375	0.026	1.398	1.7	0.9	1.399	1.7	0.9
2CA-2C	1.520	0.018	1.523	0.2	0.2	1.523	0.2	0.2
2C-3N	1.298	0.031	1.361	4.9	2.0	1.360	4.8	2.0
3N-3CA	1.460	0.014	1.462	0.1	0.1	1.463	0.2	0.2
3CA-3C	1.514	0.023	1.518	0.3	0.2	1.518	0.3	0.2
3C-4N	1.261	0.011	1.311	4.0	4.5	1.310	3.9	4.5
4N-4CA	1.471	0.015	1.487	1.1	1.1	1.487	1.1	1.1
1C-S	1.712	0.024	1.733	1.2	0.9	1.733	1.2	0.9
S-2CB	1.697	0.027	1.741	2.6	1.6	1.740	2.5	1.6
2CB-2CA	1.352	0.038	1.383	2.3	0.8	1.383	2.3	0.8
2C=O	1.228	0.028	1.241	1.1	0.5	1.242	1.1	0.5
3C-O	1.362	0.013	1.332	-2.2	2.3	1.333	-2.1	2.2
O-4CB	1.451	0.016	1.490	2.7	2.4	1.488	2.5	2.3
4CB-4CA	1.550	0.020	1.565	1.0	0.7	1.565	1.0	0.7
2N-Cu	2.077	0.015	1.981	-4.6	6.4	1.995	-3.9	5.5
3N-Cu	1.941	0.016	1.889	-2.7	3.3	1.897	-2.3	2.7
4N-Cu	2.019	0.007	1.956	-3.1	9.0	1.969	-2.5	7.1
$Cu-L_1$	2.739	0.037	2.195	-19.9	14.7	2.174	-20.6	15.3
$Cu-L_2$	1.920	0.014	2.127	10.8	14.8	2.098	9.3	12.7
N_T-N_A	2.613	0.021	2.579	-1.3	1.6	2.581	-1.2	1.5
N_A-N_O	2.585	0.032	2.547	-1.5	1.2	2.550	-1.4	1.1
N_T-N_O	4.046	0.019	3.904	-3.5	7.5	3.931	-2.8	6.1

Table 20. Bond lengths in Å for TAO-Cu^{2+} complexes with neutral ligands at B3LYP/SBKJC level (d-functions on main group atoms). Mean experimental values obtained by averaging over the four non-identical TAO sites in the unit cell of the dicopper ascidiacyclamide complex crystal structure.[69]



Angle	Mean (exp.)	SD (exp.)	$[Cu(TAO)(H_2O)_2]^+$	% Error	# of SDs	$[Cu(TAO)(MeOH)_2]^+$	% Error	# of SDs
1C-2N-2CA	111.6	3.6	111.5	-0.1	0.0	111.5	-0.1	0.0
2N-2CA-2C	118.9	2.4	116.7	-1.9	0.9	116.8	-1.8	0.9
2CA-2C-3N	109.3	1.7	108.3	-0.9	0.6	108.2	-1.0	0.6
2C-3N-3CA	119.7	2.8	118.9	-0.7	0.3	119.3	-0.3	0.1
3N-3CA-3C	102.5	2.0	103.0	0.5	0.3	103.1	0.6	0.3
3CA-3C-4N	122.9	0.6	121.2	-1.4	2.8	121.3	-1.3	2.7
3C-4N-4CA	107.5	1.4	107.4	-0.1	0.1	107.3	-0.2	0.1
1C-S-2CB	90.5	0.5	89.9	-0.7	1.2	90.0	-0.6	1.0
S-2CB-2CA	110.8	1.8	110.0	-0.7	0.4	109.9	-0.8	0.5
2CB-2CA-2N	114.4	1.3	114.3	-0.1	0.1	114.4	0.0	0.0
2CA-2C=O	120.5	2.0	122.7	1.8	1.1	122.5	1.7	1.0
3C-O-4CB	105.5	0.8	107.5	1.9	2.5	107.5	1.9	2.5
O-4CB-4CA	102.8	1.7	102.9	0.1	0.1	102.9	0.1	0.1
4CB-4CA-4N	102.3	2.0	104.1	1.8	0.9	104.2	1.9	1.0
2N-Cu-3N	81.1	0.2	83.6	3.1	12.5	83.0	2.3	9.5
3N-Cu-4N	81.5	0.8	83.0	1.8	1.9	82.5	1.2	1.3
2N-Cu-4N	162.1	0.2	165.1	1.9	15.0	165.4	2.0	16.5
3N-Cu-L_1	95.4	1.5	133.4	39.8	25.3	128.2	34.4	21.9
3N-Cu-L_2	160.9	13.4	138.9	-13.7	1.6	140.2	-12.9	1.5
N_T-N_A-N_O	102.2	0.8	99.2	-2.9	3.8	100.0	-2.2	2.8
N_A-N_T-N_O	39.1	0.2	40.1	2.6	5.0	39.7	1.5	3.0
N_A-N_O-N_T	38.6	0.6	40.7	5.4	3.5	40.3	4.4	2.8

Table 21. Bond angles in degrees for TAO-Cu^{2+} complexes with neutral ligands at B3LYP/SBKJC level (d-functions on main group atoms). Mean experimental values obtained by averaging over the four non-identical TAO sites in the unit cell of the dicopper ascidiacyclamide complex crystal structure.[69]

Uninegative ligands

The anionic ligands chosen for study were chloride and hydroxide, both of which are likely to be available for complexation in the natural environment of the parent peptides. These ligands were chosen for the additional reason that several of the fragmentation species observed in the previous MS studies on the copper complexes of the patellamides and closely related lissoclinamides had been calculated to contain either Cl^- or OH^-, presumably complexed to the copper centers.[70,71,104] Also, the fact that both had been implicated in different roles relating to the formation of carbonate species by the patellamide dicopper centers meant that they were therefore of considerable interest.[68,70,71] Geometry optimisations (B3LYP/SBKJC with d-functions on main group atoms) were initiated with the charged ligands in the "top" axial coordination site as this is the most accessible site in the natural complex and it was felt that this was the most likely position for the initial interaction with dissolved ions to occur.

$[Cu(TAO)(H_2O)_2Cl]$

Optimisation of the chloride-containing complex led to the loss of the lower axial water as seen in the optimisations of the neutral-ligand complexes. However, in this case the Cl^- was rotated until it lay in the equatorial position and the previously equatorial aquo ligand was in the lower axial position. This arrangement was thought to possibly be unfavourable in the full peptide complex where the lower axial site is considerably hindered by the Ile/Val side chains and should be expected to be subject to disturbances arising from thermal motions of these groups. As this aquo ligand had not dissociated from the complex on its own it was decided to investigate the effect of further coordination of water molecule by the copper. Another aquo ligand was introduced in the vacant top axial site at a distance of 2.40 Å from the copper (slightly more than the optimized distances found in the previous jobs). The structure was then re-optimized and the bottom axial aquo ligand was seen to dissociate leaving an approximately square pyramidal complex (Figure 34; optimized bond lengths are given in Table 22 and bond angles in Table 23).

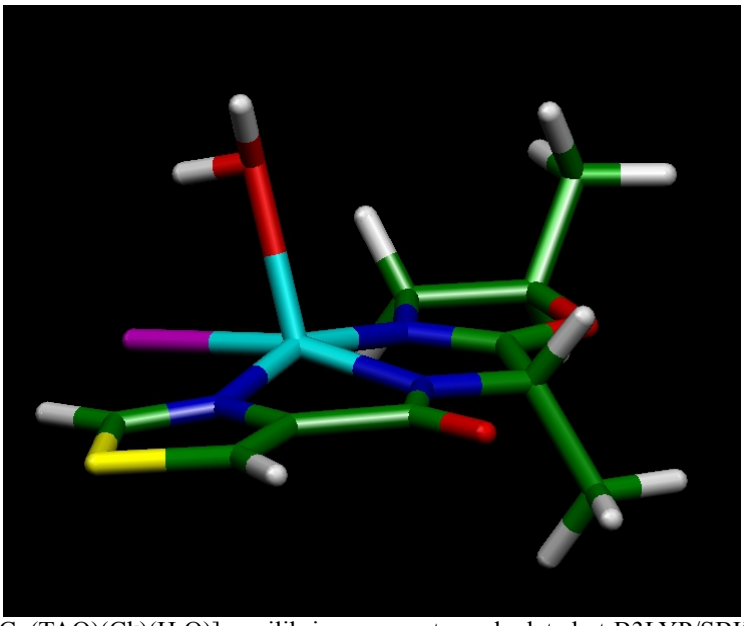


Figure 34. [Cu(TAO)(Cl^-)(H_2O)] equilibrium geometry calculated at B3LYP/SBKJC level (d-functions on main group atoms). Note distortion of Cu center upwards out of TAO plane.

[Cu(TAO)(H_2O)$_2$(OH)]

The starting point for the optimisation of the hydroxide-containing system was qualitatively identical to that used for the chloride complex. The top axial position was occupied by the OH^- ligand and the other two potential binding sites by H_2O. The initial stages of the optimisation were the same as for Cl^-, with the negatively charged hydroxide swinging down into the equatorial position and the lower axial aquo ligand dissociating. Unlike with the Cl^-, the second aquo ligand also went on to dissociate leaving a square planar copper complex formed from the TAO ligand and the OH^- (Figure 35; optimized bond lengths are given in Table 22 and bond angles in Table 23).

As with the Cl^- calculation another water molecule was introduced in the top axial position but again unlike the Cl^- complex this potential aquo ligand was ejected from the [Cu(TAO)(OH)] inner complexation sphere, leaving it hydrogen bonded to the hydroxide oxygen. As this water was no longer complexed directly to the copper it was removed and the structure re-optimized in order to be consistent with the other models. This resulted in a slightly distorted square planar complex with the hydroxide remaining in the equatorial position. The Cu-OH bond length (1.818 Å; Table 22) was noticeably shorter than that observed for the other ligands (2.416 Å (Cu-Cl), 2.127 Å (Cu-OH_2) and 2.098 Å (Cu-

OHMe)) but was similar to the equivalent Cu-O length in the ascidiacyclamide crystal structure (1.920 Å).

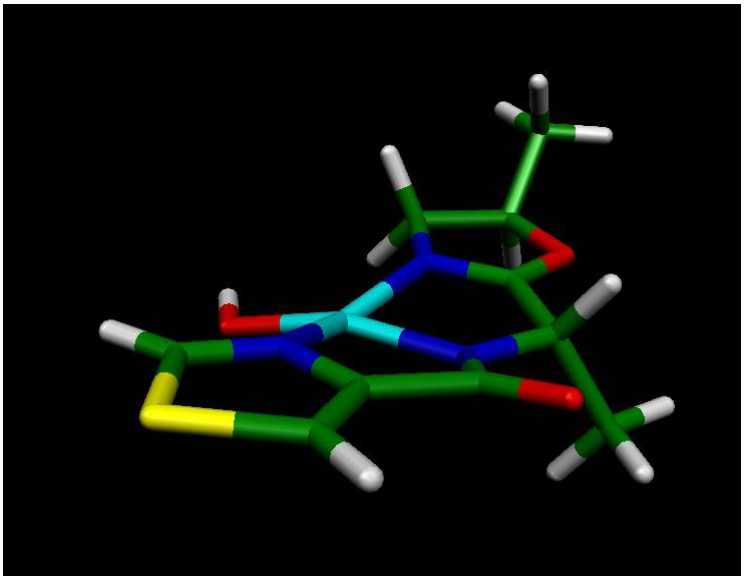
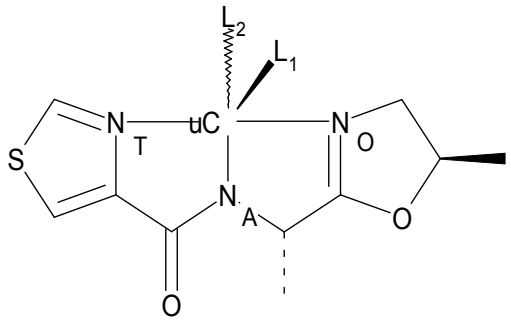
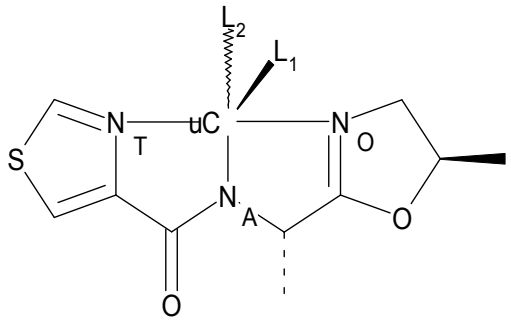


Figure 35. [Cu(TAO)(OH^-)] equilibrium geometry calculated at B3LYP/SBKJC level (d-functions on main group atoms). Note distortion of Cu center downwards out of TAO plane.



Bond	Mean (exp.)	SD (exp.)	[Cu(TAO)(Cl)(H_2O)]	% error	# of SDs	[Cu(TAO)(OH)]	% error	# of SDs
1C-2N	1.333	0.033	1.318	-1.1	0.5	1.316	-1.3	0.5
2N-2CA	1.375	0.026	1.398	1.7	0.9	1.395	1.5	0.8
2CA-2C	1.520	0.018	1.529	0.6	0.5	1.525	0.3	0.3
2C-3N	1.298	0.031	1.335	2.9	1.2	1.359	4.7	2.0
3N-3CA	1.460	0.014	1.460	0.0	0.0	1.458	-0.1	0.1
3CA-3C	1.514	0.023	1.521	0.5	0.3	1.523	0.6	0.4
3C-4N	1.261	0.011	1.300	3.1	3.5	1.303	3.3	3.8
4N-4CA	1.471	0.015	1.480	0.6	0.6	1.481	0.7	0.7
1C-S	1.712	0.024	1.740	1.6	1.2	1.742	1.8	1.3
S-2CB	1.697	0.027	1.744	2.8	1.7	1.747	2.9	1.9
2CB-2CA	1.352	0.038	1.383	2.3	0.8	1.383	2.3	0.8
2C=O	1.228	0.028	1.249	1.7	0.7	1.248	1.6	0.7
3C-O	1.362	0.013	1.344	-1.3	1.4	1.346	-1.2	1.2
O-4CB	1.451	0.016	1.486	2.4	2.2	1.484	2.3	2.1
4CB-4CA	1.550	0.020	1.566	1.0	0.8	1.566	1.0	0.8
2N-Cu	2.077	0.015	1.967	-5.3	7.3	1.985	-4.4	6.1
3N-Cu	1.941	0.016	1.908	-1.7	2.1	1.900	-2.1	2.6
4N-Cu	2.019	0.007	1.945	-3.7	10.6	1.982	-1.8	5.3
Cu-L_1	2.739	0.037	2.276	-16.9	12.5	n/a	n/a	n/a
Cu-L_2	1.920	0.014	2.416	25.8	35.4	1.818	-5.3	7.3
N_T-N_A	2.613	0.021	2.553	-2.3	2.9	2.570	-1.6	2.0
N_A-N_O	2.585	0.032	2.536	-1.9	1.5	2.570	-0.6	0.5
N_T-N_O	4.046	0.019	3.849	-4.9	10.4	3.868	-4.4	9.4

Table 22. Bond lengths in Å for TAO-Cu^{2+} complexes with charged ligands at B3LYP/SBKJC level (d-functions on main group atoms). Mean experimental values obtained by averaging over the four non-identical TAO sites in the unit cell of the dicopper ascidiacyclamide complex crystal structure.[69]



Angle	Mean (exp.)	SD (exp.)	[Cu(TAO)(Cl)(H_2O)]	% Error	# of SDs	[Cu(TAO)(OH)]	% Error	# of SDs
1C-2N-2CA	111.6	3.6	112.3	0.6	0.2	112.3	0.6	0.2
2N-2CA-2C	118.9	2.4	116.1	-2.4	1.2	115.8	-2.6	1.3
2CA-2C-3N	109.3	1.7	107.5	-1.6	1.1	108.3	-0.9	0.6
2C-3N-3CA	119.7	2.8	120.0	0.3	0.1	118.0	-1.4	0.6
3N-3CA-3C	102.5	2.0	102.7	0.2	0.1	103.6	1.1	0.5
3CA-3C-4N	122.9	0.6	121.2	-1.4	3.1	121.8	-0.9	1.8
3C-4N-4CA	107.5	1.4	108.1	0.6	0.4	107.6	0.1	0.1
1C-S-2CB	90.5	0.5	90.0	-0.6	1.0	89.9	-0.7	1.2
S-2CB-2CA	110.8	1.8	109.8	-0.9	0.6	109.7	-1.0	0.6
2CB-2CA-2N	114.4	1.3	114.1	-0.3	0.2	114.3	-0.1	0.1
2CA-2C=O	120.5	2.0	122.4	1.6	1.0	122.5	1.7	1.0
3C-O-4CB	105.5	0.8	106.7	1.1	1.5	106.9	1.3	1.8
O-4CB-4CA	102.8	1.7	103.5	0.7	0.4	103.3	0.5	0.3
4CB-4CA-4N	102.3	2.0	103.7	1.4	0.7	104.2	1.9	1.0
2N-Cu-3N	81.1	0.2	82.4	1.6	6.5	82.8	2.1	8.5
3N-Cu-4N	81.5	0.8	82.3	1.0	1.0	82.9	1.7	1.8
2N-Cu-4N	162.1	0.2	159.4	-1.7	13.5	154.3	-4.8	39.0
3N-Cu-L_1	95.4	1.5	114.8	20.3	12.9	n/a	n/a	n/a
3N-Cu-L_2	160.9	13.4	165.9	3.1	0.4	163.2	1.4	0.2
N_T-N_A-N_O	102.2	0.8	98.3	-3.8	4.9	97.8	-4.3	5.5
N_A-N_T-N_O	39.1	0.2	40.7	4.1	8.0	41.2	5.4	10.5
N_A-N_O-N_T	38.6	0.6	41.0	6.2	4.0	41.2	6.7	4.3

Table 23. Bond angles in degrees for TAO-Cu^{2+} complexes with charged ligands at B3LYP/SBKJC level (d-functions on main group atoms). Mean experimental values obtained by averaging over the four non-identical TAO sites in the unit cell of the dicopper ascidiacyclamide complex crystal structure.[69]

Discussion

The results of the density functional calculations performed here indicate that when situated within the TAO binding site found in the patellamides, Cu^{2+} displays trigonal bipyramidal coordination geometry with either H_2O or methanol as the ternary ligands. The other possible geometries for the metal in this situation (octahedral, square pyramidal or square planar) are not observed. The possibility that energetic barriers to the interconversion between these forms might exist is ruled out (at this level of theory) by the observations that both ejection of the third H_2O or methanol ligand from the primary coordination sphere of the metal and the subsequent reorganisation of the remaining ligands occurred during simple geometry optimisations. If the system was required to pass over such energy barriers, multiple minima would exist along the potential energy surface describing these processes and as a consequence, more than one (local) minimum energy structure would be found for each system.

The close similarity of the final structures of the H_2O and methanol adducts is suggestive of the possibility that the use of the latter solvent in experimental studies on patellamide-Cu^{2+} complexes may have a negligible, or at least very small, effect on the properties of the primary coordination sphere within the copper center. Explicit calculation of any such properties should, however, be made before conclusions are drawn concerning any similarity between the systems other than that observed for the coordination geometry (see next section).

In interactions involving single uninegative anionic ligands, the TAO/Cu^{2+} complex displays a strong directing effect on the equilibrium coordination geometry in that ligands of this type are preferentially complexed trans to the formally negatively charged amido nitrogen. In the full patellamide complexes containing two copper sites this behavior would mean that after initially coordinating an anionic ligand at one of the accessible top axial sites, this ligand would then displace any coordinated solvent and be rotated into a position between the coppers where it could become involved in a bridging interaction between the two metal centers. The increase of solvent lability following coordination of chloride ion has already been discussed by Burgess, 1978.[79]

It is interesting to note that changes in coordination geometry similar to those observed in

the TAO complexes were observed by Morokuma and co-workers, 1995, in an *ab-initio* Hartree-Fock study of the substitution of an aquo ligand by chloride in a square planar Pt(II) complex.[105,106] In this study a chloride ion was introduced at one of the axial sites in a (square-planar) $[Pt(NH_3)_3(H_2O)]^{2+}$ complex and the geometries and energetics of stationary points on the substitution surface calculated. Pt-N lengths and N-Pt-N angles were fixed at their experimental geometries and not allowed to vary during calculations. These workers found that the introduction of the chloride ion lead to the displacement of the equatorial water towards the other axial site. The vacated equatorial position was then occupied by the chloride and the displaced aquo ligand was left weakly coordinated to the Pt center and hydrogen-bonded to the chloride. Several saddle points were located on the potential surface for this process but otherwise the substitution occurred in a similar fashion to that seen for the TAO complexes in the present work.

The nature of the anion also appears to be important in determining the equilibrium features of the TAO/Cu^{2+} site in that the smaller of the two ligands studied (OH^-) seems to reduce the coordinating ability (or strength of interaction) of the copper with respect to uncharged ligands. This could be expected to have an effect on the reactive properties of the site and is suggestive of the possibility that the activity of the patellamide-Cu^{2+} system may be mediated by interaction with different forms of anionic ligand.

Chapter 5

Density Functional Investigation of the Reactivities of TAO/Cu^{2+} Complexes

Introduction

The reactivities of Cu^{2+} complexes with N-donor ligands have been studied previously using density functional theory by Martinez, *et al*, 2001.[101] The authors were attempting to rationalize the experimental observation that when Cu^{2+} interacts with O-donor ligands in aqueous solution, the resulting complex is stabilized if N-donor ligands are also coordinated to the metal. The complexes investigated contained variously substituted 1,10-phenanthroline (phen) ligands as the N-donors and the remaining coordination sites on the copper were occupied with aquo ligands ($[Cu(phen)(H_2O)_3]^{2+}$) or a combination of aquo ligands and doubly deprotonated oxalate (ox) ion ($[Cu(phen)(ox)(H_2O)]$). The hydrated Cu^{2+} ($[Cu(H_2O)_5]^{2+}$) was included for comparison. Density functional calculations were performed with the B3LYP functional and a 3-21G(d,p) basis set (single-point calculations at the B3LYP/3-21G(d,p) geometries using the 3-21G and 6-311G(d,p) sets were performed in order to check for basis set effects but the results were found to be qualitatively the same).

The authors found an inverse correlation between the electron affinity (π-acceptor properties) of the complexes with aquo ligands and the pK_a values of the differently substituted phenanthrolines. This effect was also found in the oxalate-containing complexes. The electron affinities of the phenanthroline-containing complexes were seen to be larger than that observed for the hydrated metal, suggesting that the phenanthroline complexes should react to form more stable complexes with negatively charged oxygens than the aquo complexes in line with the experimental observations.

These results were then re-interpreted in terms of the density functional theory of chemical reactivity (see next section). Sigel *et al.*[107] had previously suggested that an increase in the hardness of Cu^{2+} on complexing with N-donor ligands was responsible for the increased stability of mixed N/O-donor complexes. The global hardness[61] and local softness[108] (characterized by the corresponding value of the Fukui function[109]) were calculated and the trends in these properties compared with the experimental data. In contrast to the suggestion of Sigel *et al.*,[107] $[Cu(phen)(H_2O)_3]^{2+}$ was found to be softer than $[Cu(H_2O)_5]^{2+}$. The Fukui function describing susceptibility to nucleophilic attack correlated well with the values of the stability constants for the complexes. From this it was concluded that the

local chemical softness was a good descriptor of reactivity in these mixed Cu^{2+} complexes but the use of the global hardness (equivalent to Pearson's original definition in the theory of hard and soft acids and bases) was inappropriate for this purpose.

Density Functional Theory of Chemical Reactivity

The Fukui function of Parr and Yang, 1984,[109] describes the differing reactivity at different points in space for a given chemical species by evaluating the change in electron density, ρ, with the electron number, N

$$f(\vec{r}) = \left(\frac{\partial \rho(\vec{r})}{\partial N} \right)_{v(\vec{r})} \tag{5.1}$$

with large values of $f(\vec{r})$ indicating areas where the molecule or ion is chemically soft. The relationship between $f(\vec{r})$ and the chemical softness is given by

$$\sigma(\vec{r}) = f(\vec{r})S \tag{5.2}$$

where $\sigma(\vec{r})$ and S are the local and global softnesses, respectively.[108] The latter is obtained by taking the reciprocal of the global chemical hardness, η, discussed earlier in Chapter 3. It can be seen from equation 5.1 above that $f(\vec{r})$ contains all the information on site-specific reactivity contained in $\sigma(\vec{r})$ and consequently the Fukui function alone may be used as an indicator of chemical reactivity at a given (local) resolution defined by the coordinates $\vec{r}$.

This approach to chemical reactivity is in fact the generalization of the frontier molecular orbital (FMO) method within density functional theory. In FMO, the highest-occupied and lowest-unoccupied molecular orbitals (HOMO and LUMO, respectively) are taken as being most important in determining the outcome of a (covalent bond-forming) reaction. The success of FMO can be shown to be due to the fact that the HOMO and LUMO reproduce (approximately) features in the total charge density and it is these features which in fact control reactivity.[40] Whereas FMO neglects contributions from the remainder of the orbitals in the system it can be seen that the Fukui function implicitly contains any effects due to these orbitals which might be of importance in deciding the course of a

reaction.

A finite difference approximation to the differential form of the Fukui function given above was introduced by Yang and Mortier, 1986.[110] This approach makes use of condensed atomic electronic populations, n_i, generated by the Mulliken orbital partitioning scheme and allows the separation of $f(\vec{r})$ into separate functions for nucleophilic, $f^+(\vec{r})$, electrophilic, $f^-(\vec{r})$, and neutral (radical) attack, $f^0(\vec{r})$. Thus the Fukui function for the copper center in the complexes under study may be written

$$f^+_{Cu} = n_{Cu}[N_{TAO}+1] - n_{Cu}[N_{TAO}] \tag{5.3}$$

$$f^-_{Cu} = n_{Cu}[N_{TAO}] - n_{Cu}[N_{TAO}-1] \tag{5.4}$$

$$f^0_{Cu} = \frac{1}{2}\left(n_{Cu}[N_{TAO}+1] - n_{Cu}[N_{TAO}-1]\right) \tag{5.5}$$

where $n_{Cu}[N_{TAO}]$, $n_{Cu}[N_{TAO}+1]$ and $n_{Cu}[N_{TAO}-1]$ are the Mulliken populations of the Cu atom in the TAO complex with N electrons and after one-electron reduction or oxidation respectively.

Method

All calculations were carried out at the B3LYP/6-31G(d,p) level at the N-electron geometries in line with the requirement that $\partial\rho(\vec{r})/\partial N$ be evaluated at constant external potential (in this case the B3LYP geometries obtained with the polarized SBKJC basis were used). In addition to the Fukui function for nucleophilic attack calculated by Martinez, *et al*., those relating to electrophilic and radical attack were also calculated. The same protocol was used to calculate geometries and values of f_{Cu} for the hydrated Cu^{2+} species, $[Cu(H_2O)_5]^{2+}$, and the aquo-Cu^{2+} complex of 1,10-phenanthroline, $[Cu(phen)(H_2O)_3]^{2+}$, for comparison both with their literature values[101] and with the values obtained in this study for the TAO complexes to provide reference points for use in the assessment the effects of complexation on the Cu^{2+} ion.

Geometries of $[Cu(H_2O)_5]^{2+}$ and $[Cu(phen)(H_2O)_3]^{2+}$

Selected geometrical parameters for both complexes were compared with those obtained from the supporting material accompanying reference 101 are shown in Table 24. In the $[Cu(H_2O)_5]^{2+}$ complex the bond lengths obtained in this work (B3LYP/SBKJC with d-functions on main group atoms) deviated from the literature values by 1.0% (0.02 Å, axial Cu-O) and 3.1% (0.06 Å, equatorial Cu-O). The calculation on the $[Cu(phen)(H_2O)_3]^{2+}$ complex gave bond lengths that were 2.5% (0.05 Å, Cu-O (equatorial)) and 1.6% (0.03 Å, Cu-N) greater than the equivalent lengths obtained from the B3LYP/3-21G(d,p) calculations. The N-Cu-N angle deviated from the literature value by -0.4% (0.3 Å). The overall geometries displayed distorted trigonal bipyramidal symmetry (Figure 36).

$[Cu(H_2O)_5]^{2+}$

Bond	A	B	Deviation (Å)	Deviation (Å)
Cu-OH_2 (ax.)	1.92	1.94	0.02	1.0
Cu-OH_2 (eq.)	1.94	2.00	0.06	3.1

$[Cu(phen)(H_2O)_3]^{2+}$

Bond	A	B	Deviation (Å)	Deviation (Å)	Angle	A	B	Deviation (°)	Deviation (°)
Cu-OH_2 (eq.)	1.97	2.02	0.05	2.5	1N-Cu-2N	86.1	85.8	0.3	0.4
Cu-N	1.92	1.95	0.03	1.6					

Table 24. Comparison of geometrical parameters for $[Cu(H_2O)_5]^{2+}$ and $[Cu(phen)(H_2O)_3]^{2+}$ calculated with B3LYP functional. A. 3-21G(d,p) values obtained from literature (supporting information accompanying publication).[101] B. Values from this work (SBKJC basis with d-functions on main group atoms). Bond lengths in Å and angles in degrees.

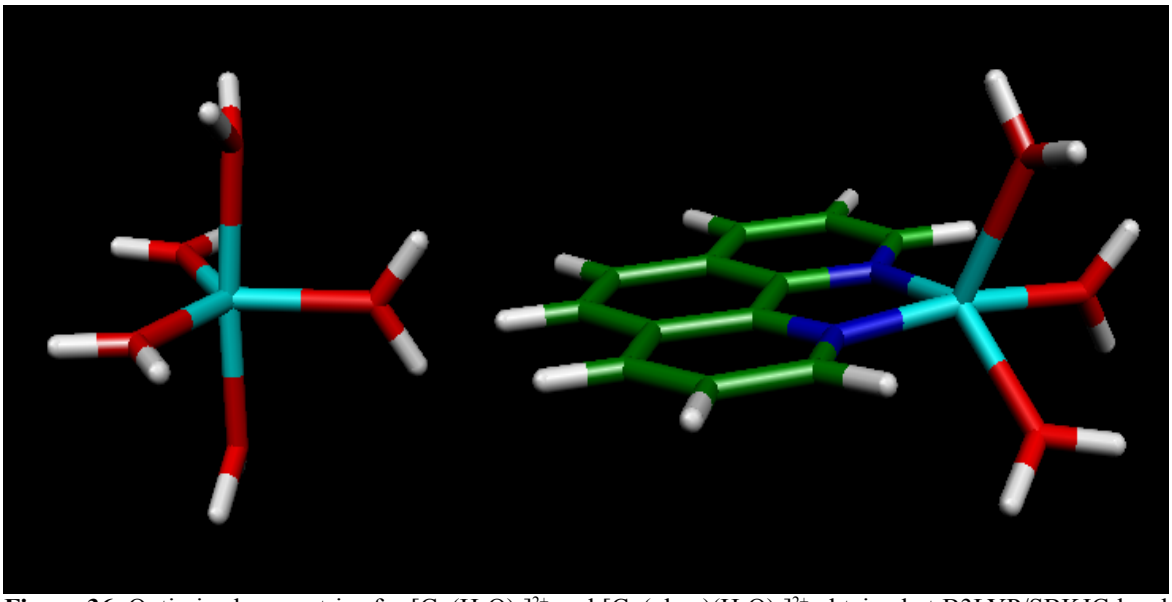


Figure 36. Optimized geometries for $[Cu(H_2O)_5]^{2+}$ and $[Cu(phen)(H_2O)_3]^{2+}$ obtained at B3LYP/SBKJC level (d-functions on main group atoms).

Values of density functional reactivity descriptors for copper centers

The results of the reactivity calculations are summarized in Table 25 along with the available literature values for $[Cu(H_2O)_5]^{2+}$ and $[Cu(phen)(H_2O)_3]^{2+}$. The global and local properties calculated in this work are compared graphically in Figures 37 & 38.

		Global Properties			Local Properties		
Complex	q_{Cu}^{M}	I	A	η	f_{Cu}^{+}	f_{Cu}^{-}	f_{Cu}^{0}
$[Cu(H_2O)_5]^{2+}$	1.181	19.952	9.050	5.451	0.532	0.172	0.352
*$[Cu(H_2O)_5]^{2+}$	-	18.430	6.040	6.190	0.600	-	-
$[Cu(phen)(H_2O)_3]^{2+}$	1.127	14.708	8.033	3.338	0.374	0.047	0.210
*$[Cu(phen)(H_2O)_3]^{2+}$	-	14.980	6.140	4.420	0.270	-	-
$[Cu(TAO)(H_2O)_2]^{+}$	1.026	10.678	3.676	3.501	0.356	0.101	0.228
$[Cu(TAO)(MeOH)_2]^{+}$	1.054	10.497	3.724	3.386	0.341	0.092	0.216
[Cu(TAO)(H_2O)Cl]	0.932	7.523	0.048	3.738	0.264	0.061	0.162
[Cu(TAO)(OH)]	0.930	7.117	0.381	3.368	0.360	0.151	0.255

Table 25. Summary of values for reactivity descriptors calculated at B3LYP/6-31G(d,p) level. Mulliken partial charge (q_{Cu}^{M}), ionization potential (I, eV), electron affinity (A, eV), global hardness (η) and condensed Fukui functions (f_{Cu}). *Values from Martinez, *et al.*[101]

Comparison with literature values

Comparing the values obtained here for the $[Cu(H_2O)_5]^{2+}$ and $[Cu(phen)(H_2O)_3]^{2+}$ complexes with those of Martinez, *et al*. it can be seen that the use of the larger 6-31G(d,p) basis gave some differences in the calculated properties. The values of I and A

were both increased for the hydrated copper whereas in the $[Cu(phen)(H_2O)_3]^{2+}$ complex the changes were negative in I and positive in A. These changes are most probably due to the improved description of the electronic structure of the valence region of the complexes provided by the larger basis set (and consequently of the σ- and π-interactions within the complex). This was considered to be an important observation as the extent of unsaturation in the TAO ligand meant that the improved treatment of these regions by the larger, more flexible, basis set would be important in the study of its complexes. Furthermore, the π-systems in the TAO ligand contained a larger number of elements (C, N, O and S) than the π-systems of the phen ligand (C and N) meaning that the polarizability of the basis set would be even more important.

A reduction in the HOMO-LUMO energy spacing was implied by the changes in the values of I and A and this could be observed in the lowering of the hardness values for both complexes (see section on theory, above). The difference between the literature values for the Fukui function and those calculated in this work could not be rationalized properly because the Mulliken charges used in the literature calculations were not reported.

Properties of TAO site with neutral ligands

$[Cu(TAO)(H_2O)_2]^+$

The systemic formal charge was changed from +2 to +1 on going from the reference complexes to those containing the TAO ligand. However, the Mulliken partial charge on the copper in the $[Cu(TAO)(H_2O)_2]^+$ system was only lowered by 0.155 electrons and 0.101 electrons when compared with $[Cu(H_2O)_5]^{2+}$ and $[Cu(phen)(H_2O)_3]^{2+}$, respectively. The values for the condensed Fukui functions showed that the nucleophilic and radical susceptibilities of the TAO copper were similar in magnitude to those of $[Cu(phen)(H_2O)_3]^{2+}$. The electrophilic susceptibility was intermediate between the values for the two reference systems.

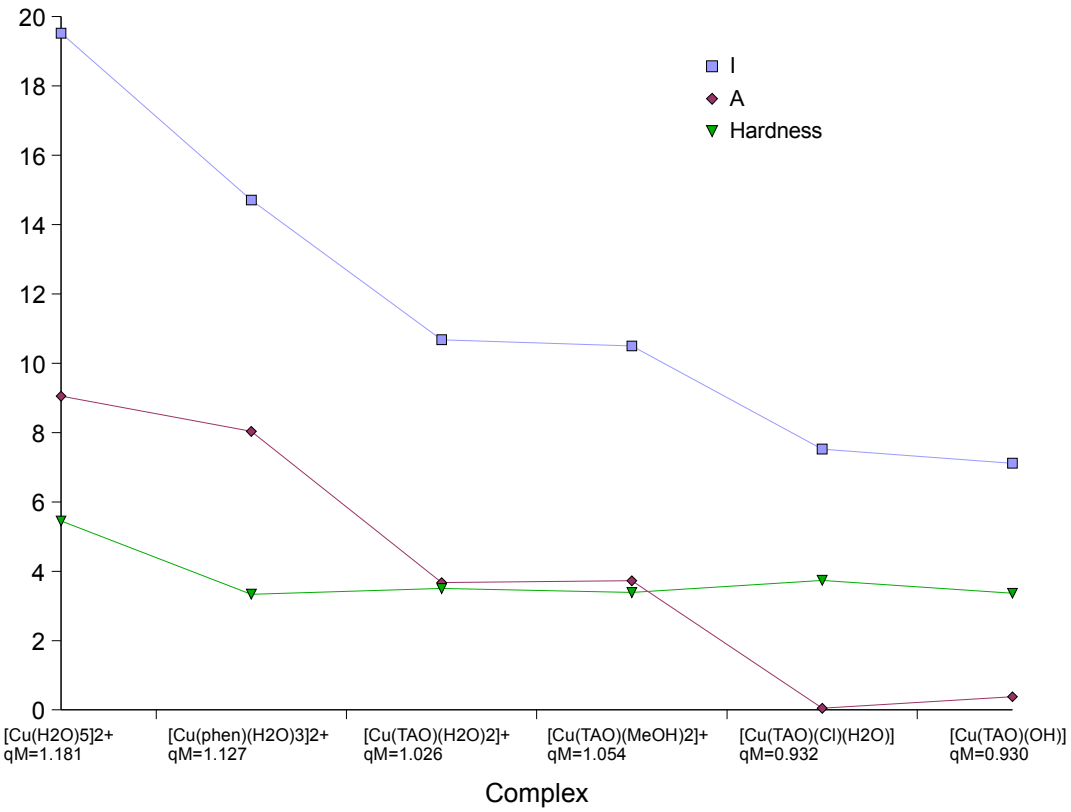


Figure 37. Global properties calculated at the B3LYP/6-31G(d,p) level. The Mulliken partial charges (qM) found for copper are shown on the x-axis. Ionization potential (I) and electron affinity (A) in eV, hardness in dimensionless units. Data points joined by lines for increased clarity.

$[Cu(TAO)(MeOH)_2]^+$

The calculated properties of the methanolic TAO complex were almost identical to those o f $[Cu(TAO)(H_2O)_2]^+$. A slight increase in the partial charge on the copper center with respect to the hydrated TAO complex was accompanied by low magnitude reductions in all three Fukui functions.

Properties of TAO site with uninegative ligands

$[Cu(TAO)(H_2O)(Cl)]$

Substitution of one of the aquo ligands in $[Cu(TAO)(H_2O)_2]^+$ by chloride ion resulted in the neutralization of the formal charge but the partial charge on the copper was again only reduced by a similar amount to that observed on going from $[Cu(H_2O)_5]^{2+}$ to $[Cu(TAO)(H_2O)_2]^+$ (0.094 electrons).

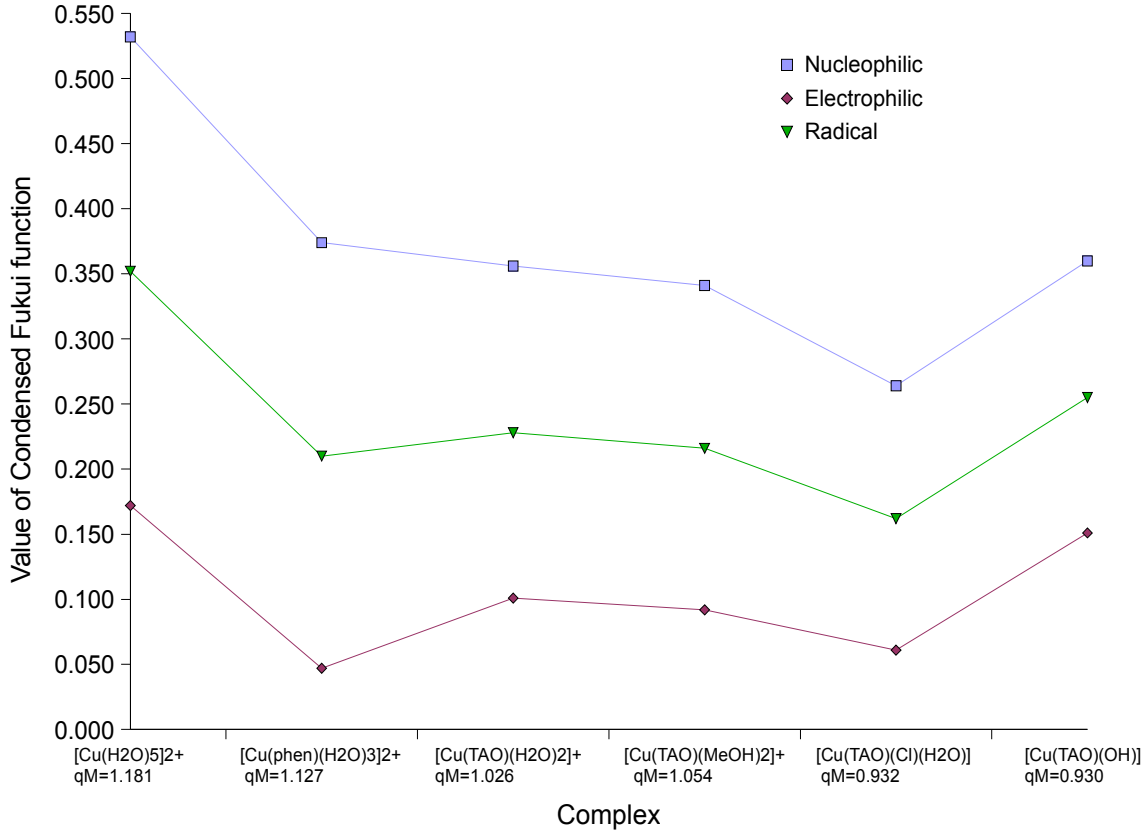


Figure 38. Local properties (condensed Fukui functions) calculated at the B3LYP/6-31G(d,p) level. The Mulliken partial charges (qM) found for copper are shown on the x-axis. Data points joined by lines for increased clarity.

The ionization potential was reduced by approximately 25% (in line with the change on going from $[Cu(phen)(H_2O)_3]^{2+}$ to $[Cu(TAO)(H_2O)_2]^{+}$). A slight increase in the hardness was accompanied by the reduction of the electrophilic susceptibility to approximately half of its value in $[Cu(TAO)(H_2O)_2]^{+}$. The electron affinity was reduced almost to zero and this was reflected in decreases in the nucleophilic and radical susceptibilities by approximately one third.

[Cu(TAO)(OH)]

The partial charge on the copper of the [Cu(TAO)(OH)] system was 0.002 electrons less than that found in $[Cu(TAO)(H_2O)(Cl)]$. The global properties calculated showed almost identical deviations from the $[Cu(TAO)(H_2O)_2]^{+}$ values to those of $[Cu(TAO)(H_2O)(Cl)]$. Unlike $[Cu(TAO)(H_2O)(Cl)]$ the reactive susceptibilities were found to be higher than in

$[Cu(TAO)(H_2O)_2]^+$. The value for electrophilic susceptibility was closer to that found in $[Cu(H_2O)_5]^{2+}$ than the values for any of the other complexes studied and was 50% larger than in $[Cu(TAO)(H_2O)_2]^+$.

Discussion

The results of the present study into the reactivities of ternary Cu^{2+} complexes with the TAO binding site found in the patellamides and neutral O-donor ligands indicate a similar susceptibility to nucleophilic attack to that found in the 1,10-phenanthroline complex $[Cu(phen)(H_2O)_3]^{2+}$. The reduction of this value in both the TAO and phenanthroline complexes relative to $[Cu(H_2O)_5]^{2+}$ is indicative of the effects of the electronic environment created by the organic ligands. This is to be expected as both the TAO and phen ligands should display π-acceptor activity due to the level of unsaturation within their structures. H_2O is a simple σ-donor ligand and therefore does not enter into this type of interaction. As a result, both organic ligands would be expected to redistribute and stabilize the increased negative charge associated with the approach of an electron-rich species to a greater degree than aquo ligand alone.

Similar conclusions can be drawn from the data concerning radical susceptibility. Again, the values for the TAO-containing complexes are similar to that of $[Cu(phen)(H_2O)_3]^{2+}$ and considerably reduced compared with $[Cu(H_2O)_5]^{2+}$. A slight increase in the value for $[Cu(TAO)(H_2O)_2]^{2+}$ over those for $[Cu(phen)(H_2O)_3]^{2+}$ and $[Cu(TAO)(H_2O)_2]^{2+}$ indicates a direct correlation between radical susceptibility and global hardness. The charge-delocalisation effects discussed for the case of nucleophilic attack are not expected to have the same importance for initial interactions with radicals (this cannot of course be concluded for the later stages of such a reaction without performing the relevant calculations).

The electrophilic susceptibility displayed by the TAO-containing complexes is considerably larger than that of $[Cu(phen)(H_2O)_3]^{2+}$ reflecting the large decrease in the ionization potential in these systems. It is possible that both of these effects might be attributed to the presence of a sulfur atom in the TAO ligand. Sulfur is an important element in the ligand environments of many biological copper-based redox systems and its

involvement in the 10-electron aromatic thiazoleamide moiety of the TAO ligand makes it likely that it plays an important role in regulating the reactive properties of the complex.
The substitution of the aquo ligands in $[Cu(TAO)(H_2O)_2]^+$ by methanol caused only very minor changes in the equilibrium properties of the complex at both the global and local levels. This is an important finding for future experimental studies on patellamide-Cu^{2+} complexes as problems with water-solubility mean that methanol is commonly used as a substitute. The results of this theoretical study support the continued use of this solvent for this purpose.

The substitution of uninegative charged ligands into the $[Cu(TAO)(H_2O)_2]^+$ complex can be seen to cause considerable changes in the reactivities of the copper center. Chloride causes all three reactive susceptibilities to fall below the levels seen in the $[Cu(TAO)(H_2O)_2]^+$ system whilst the effect of hydroxide ion is to cause these values to rise above those of $[Cu(TAO)(H_2O)_2]^+$. This is quite surprising as the global properties of the two neutral complexes are very similar (as is the partial charge on the copper). The depression of nucleophilic susceptibility by chloride ion may be due to donation of electron density from its 3p region to the π-acceptor system of the TAO ligand. The effect of such donation would be to reduce the ability of the ligand to accommodate the increase in electron density occurring during nucleophilic attack. Conversely, increased delocalisation of the unpaired 3d-electron from the copper into the empty 3d region of the chloride ligand might reasonably be expected to reduce the metal's susceptibility to electrophililic or radical attack, in line with the results of this study.

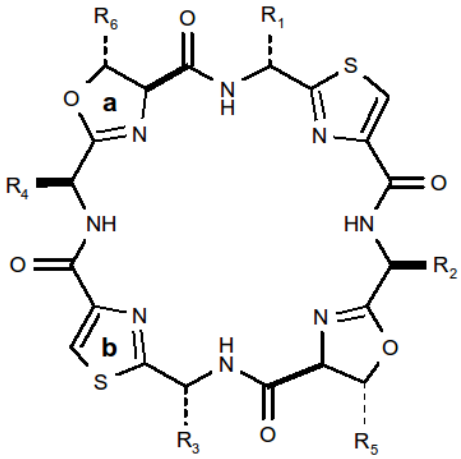
Because of the lack of either filled 3p or empty 3d orbitals on the hydroxide ligand, neither of the effects mentioned above are possible in [Cu(TAO)(OH)]. As a result, a small (compared to the chloride case) reduction of the nucleophilic susceptibility might be expected along with slight increases in the electrophilic/radical susceptibility due simply to the reduction in the positive partial charge on the copper. The large increase in all three local reactivities does not fit in with this suggestion but it seems feasible that, in combination with the charge-effects of the hydroxide, the greater unsaturation in the primary coordination sphere of the copper caused by loss of the axial aquo ligand may be responsible. If this hypothesis is correct, the fact that the axial aquo ligand was ejected

from the primary coordination sphere of the copper during the geometry optimizations (despite the fact that H_2O is itself a weak nucleophile) suggests that [Cu(TAO)(OH)] displays selectivity with respect to the nature of the nucleophilic species.

Concluding Remarks and Future Work

Introduction

The main aim of the present work was to investigate the conformational preferences of the patellamides (**1**-**8**) and the coordination environment of their Cu^{2+} complexes in the absence of complicating factors such as solvation and the effect of the crystal environment. This was achieved through the use of density functional theory (DFT) to obtain equilibrium geometries and energies for a subset of the patellamide family and to a model of the TAO binding site (Figure 40) located within the macrocyclic cavity of these peptides. Density functional theory was also applied to the elucidation/rationalization of some of the properties of these compounds, yielding several interesting insights into the possibilities for non-covalent interactions in the patellamides and the varying reactivity of their Cu^{2+} complexes.



Patellamides:

1 Ascidiacyclamide R_1=R_3=D-Val R_2=R_4=L-Ile
2 Patellamide A R_1=R_3=D-Val R_2=R_4=L-Ile
3 Patellamide B R_1=D-Phe R_2=L-Leu R_3=D-Ala R_4=L-Ile
4 Patellamide C R_1=D-Phe R_2=L-Val R_3=D-Ala R_4=L-Ile
5 Patellamide D R_1=D-Phe R_2=R_4=L-Ile R_3=D-Ala
6 Patellamide E R_1=D-Phe R_2=R_4=L-Val R_3=D-Ile
7 Patellamide F R_1=D-Phe R_2=R_4=Val R_3=Ile
8 Patellamide G R_1=D-Phe R_2=L-Ile R_3=D-Ala R_4=L-Ile

Note: R_5=R_6=CH_3 for all compounds except patellamide A (**2**) where R_5=CH_3 and R_6=H

Figure 39. Schematic diagram of the molecular structure of the patellamides (**1**-**8**). Five-membered ring systems: **a**=oxazoline, **b**=thiazole.

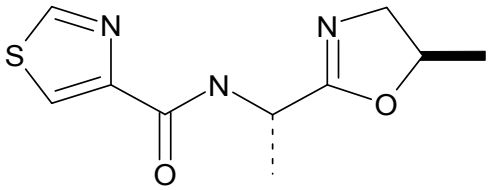


Figure 40. Model structure used to represent the TAO binding site found in the patellamides.

Equilibrium geometries and conformational energetics of the patellamides

The combination of the energetic and geometric data obtained from DFT calculations at the B3LYP/6-31G(d,p) level (performed with the assistance of Dr. Martyn F. Guest at the

CLRC Daresbury Laboratories) provides an insight into the sources of conformational preferences displayed by the patellamides. These data are in agreement with, and help to rationalize, the results of the empirical studies performed to date. The patellamides studied here display positive energy changes on going from the open to closed conformations ($\Delta E_{open \rightarrow closed} > 0$; Table 26). Removal of a single methyl group from ascidiacyclamide (ASC; **1**) at one of the oxazoline rings to give patellamide A (PatA; **2**) leads to changes in the β-turn forming properties of the molecule and a reduction in $\Delta E_{open \rightarrow closed}$ of ~25%. The larger substitutions observed in patellamides C (PatC; **4**) and D (PatD; **5**) alter the local structure around both the oxazoline and thiazole rings in these compounds, allowing these molecules to close more completely and this substitution pattern is accompanied by a reduction in $\Delta E_{open \rightarrow closed}$ of ~75% compared with the C_2 symmetric ascidiacyclamide.

Peptide (preferred conformation)	B3LYP/6-31G(d,p)	MM2*	AMBER*
ASC (open)	21.40	-59.81	-57.00
PatA (open)	16.40	-60.74	-59.98
PatC (closed)	5.80	-67.47	-52.09
PatD (closed)	3.60	-72.28	-47.33

Table 26. Conformational energy differences (open→closed) for ASC (**1**), PatA (**2**), PatC (**4**) and PatD (**5**).

From consideration of these findings it seems that there are two separable mechanisms operating in the patellamides that might allow for alterations in the ability of the structures to fold into the closed conformation. The first, desymmetrization of the oxazoline rings causes changes in the β-turn forming properties of the molecule and the second, substitution at R_1 and R_3, has its effect by explicitly altering the geometry outwith the β-turns such that the molecule closes more completely than the C_2 symmetric ascidiacyclamide (reflected in the reductions in the lengths of the hydrogen bonds in the closed conformation (Table 27) and the transannular separation of the thiazole rings (Table 28)). The involvement of the oxazoline rings in the folding of the patellamides was demonstrated in a previous publication (a preliminary study performed as part of the present work) by the author and co-workers (see Appendix – Milne *et al.*, 2002) using MD simulations and goes some way to support these conclusions.

Patellamide	Carbonyl donor oxygen		Oxazoline donor oxygen	
	1NH-6O	5NH-2O	3NH-7O	7NH-3O
ASC	2.406	2.406	2.491	2.491
PatA	-0.017	-0.010	-0.025	-0.106
PatC	-0.131	-0.167	-0.170	-0.107
PatD	-0.043	-0.075	-0.015	-0.085

Table 27. Hydrogen bonding in patellamide closed conformation (B3LYP/6-31G(d,p) geometries). Distances are shown explicitly for ASC and deviations from these values are shown for PatA/C/D. Measurements in Å.

Patellamide	Transannular separation of thiazole rings (Å)
ASC	4.34
PatA	4.33
PatC	4.09
PatD	4.04

Table 28. Transannular separation of thiazole rings in patellamide closed conformers (B3LYP/6-31G(d,p) geometries). Values correspond to measurements made between centers of thiazole rings.

Comparing the DFT energies with those obtained from molecular mechanics calculations with the MM2* and AMBER* force fields have highlighted an important deficiency with respect to the use of this approach to study the conformational behaviour of the patellamides. MM2* performed reasonably well in predicting the correct conformational trend but failed to provide conformational energies even close to the DFT values. AMBER* failed to predict either the trend or the correct energy changes connecting the open and closed conformers in the patellamides studied. This is an important finding both for future studies on cyclic peptides of this type and for the interpretation of data already existing in the literature as incorrect assumptions could quite easily be drawn from the results of this type of calculation if this source of error was to remain unnoticed. In any such future studies on the patellamides in the gas phase the findings of the present work should be borne in mind.

Charge-transfer properties of the patellamides

The net effects of the varying substitution pattern in the patellamides studied here are also evident in the discussion of their maximal intermolecular charge-transfer capabilities (Chapter 3). The effect of amino acid substitutions in the patellamides on the conformation-dependent change in the charge transfer energy $(\Delta\Delta E_{MAX}^{CT})$ is shown in Table 29 along with the results of including this as a correction to the ground state

B3LYP/6-31G(d,p) energies (ΔE^{DFT}). ΔE^{CT}_{MAX} was calculated using the following finite difference expressions obtained from density functional theory:[67]

$$\Delta E^{CT}_{MAX} = -\frac{I+A}{2}\Delta N_{MAX} + \frac{I-A}{2}(\Delta N_{MAX})^2$$

$$\Delta N_{MAX} = \frac{(I+A)}{2(I-A)}$$

Peptide	$\Delta\Delta E^{CT}_{MAX}$	ΔE^{DFT}	ΔE^{TOTAL}
ASC	1.7	21.4	23.1
PatA	0.04	16.4	16.4
PatC	-4.5	5.8	1.3
PatD	-2.4	3.6	1.2

Table 29. Effect of conformational change on global maximum ground state charge-transfer energies of the patellamides $(\Delta\Delta E^{CT}_{MAX})$. All values given in kJ mol^{-1}. ΔE^{TOTAL} is the result of including $\Delta\Delta E^{CT}_{MAX}$ as a correction to ΔE^{DFT} (obtained from B3LYP/6-31G(d,p) calculations). $\Delta\Delta E^{CT}_{MAX}$ for PatA is shown to two decimal places in order that its non-zero value can be seen.

Whilst this treatment is not particularly rigorous (*e.g.* the estimation of the electron affinities of the patellamides (used in the calculation of $\Delta\Delta E^{CT}_{MAX}$) from the energies of their lowest unoccupied molecular orbitals is questionable as it is known that this energy is, to some extent, dependent on the basis set used in the calculation of the orbitals),[40] it is felt that this topic is worthy of further study as the effects of partial charge-transfer are likely to be of importance in altering the properties of biological molecules such as the patellamides.

A more reliable method for obtaining values of both the ionization potential, I, and electron affinity, A, would be to perform three single point energy calculations on the equilibrium patellamide structures with global charges of -1, 0 and +1 and then subtract these energies to obtain the 'vertical' values of I and A.[101] In the author's experience, the estimate of I obtained from the HOMO energy is usually in reasonably good agreement with the vertical ionization potential (the former is perhaps slightly smaller in magnitude)

but the corresponding estimate of A obtained from LUMO energy is generally significantly smaller than its explicitly calculated value. Calculating I and A in this manner is expected to bring about changes in $\Delta\Delta E_{MAX}^{CT}$ and it is felt that this should be investigated before further conclusions are drawn concerning the effects of intermolecular charge-transfer in the patellamides studied here.

Intramolecular charge-transfer in the patellamides was also discussed within the framework of density functional theory and it was concluded that energy lowering due to transannular charge-transfer between the thiazole rings in the closed conformation of the patellamides as suggested by previous workers[24,25] may be partially responsible for the preference for the closed conformation displayed by the more highly substituted and asymmetric patellamides. In the closed conformation the thiazole rings are brought close together in the correct orientation for π-stacking and the differences in the transannular separation for the different patellamides are observed (Table 28) with the phenylalanine-containing PatC and PatD displaying the closest stacking of these rings. This discussion was, however, qualitative and further calculations will be required if this hypothesis is to be tested theoretically.

One approach that might aid in this is the natural bond orbital method[40] in which an energy calculation is performed using the full set of ('occupied' and 'Rydberg') natural orbitals followed by a similar calculation in which the Rydberg orbitals have been omitted and the occupation numbers of the occupied orbitals are fixed at integer values. The difference between the two energies will give an estimate of the effects of partial intramolecular charge-transfer in the species in question.

Conformation-dependent alterations in the dipole moments of the patellamides

The dipole moments for ASC, PatA, PatC and PatD were calculated (Chapter 3) at the B3LYP/6-31G(d,p) level at the equilibrium geometries obtained earlier. The magnitudes of the dipole moments (Table 30) show that in all four patellamides the dipole moment is reduced on folding into the closed conformation. In the open conformation the magnitudes of the dipole moments are smaller in the phenylalanine-substituted PatC and PatD than those in ASC or PatA. The degree of change of the dipole moment is different for the

differently substituted peptides, with PatA displaying the largest reduction on folding.

Peptide	"open"	"closed"	Δ
ASC	3.04	1.79	-1.25 (-41.2%)
PatA	3.24	1.54	-1.70 (-52.5%)
PatC	2.92	2.16	-0.72 (-26.0%)
PatD	2.80	1.79	-1.01 (-36.1%)

Table 30. Conformational dependance of patellamide dipole moments calculated at the B3LYP/6-31G(d,p) level. Values in Debye units.

Both PatC and PatD showed smaller changes than the more symmetrical ASC and PatA. This is not the result that might be expected from consideration of the previous observation that polar solvents lead to a favouring of the closed conformation, suggesting that the gas phase dipoles given here may undergo considerable alteration upon solvation. Alternatively, it may be that the dipole moment does not play a dominant part in determining the solution conformation of these molecules.

Plotting the dipole moments with respect to the patellamide structures (ASC is given as an example in Figure 41) showed that, in addition to undergoing a reduction in magnitude the dipoles are re-oriented through 180º. In ASC the dipole moment is aligned along the C_2 symmetry axis of the molecule. The effects of the desymmetrization in the other patellamides were evident in the deviation of the dipole arrow from the position of this axis in ASC.

It was speculated that the inversion of the dipole observed in the closed conformation of the patellamides may have consequences for the interactions of these peptides with large biomolecular assemblies in that reversal of the direction and alteration in the magnitude of the dipole moment could act to mediate the strength and type of the interaction and alter the properties of the complex. In the relatively simple case of interaction with a negatively charged biological membrane surface (assuming that no binding site specific for the patellamide structure is involved), the peptide would be expected to orient itself so that the positive end of the dipole came into contact with the membrane surface. Whilst this arrangement would be stable, if the patellamide were to undergo conformational change the dipole would be reoriented so that the negative end was now pointed at the negatively

charged membrane and consequently the interaction would become unfavourable, leading possibly to complete dissociation of the patellamide from the surface. Inclusion of the conformational effects of intermolecular charge-transfer mentioned above would seem to make this hypothetical model even more plausible.

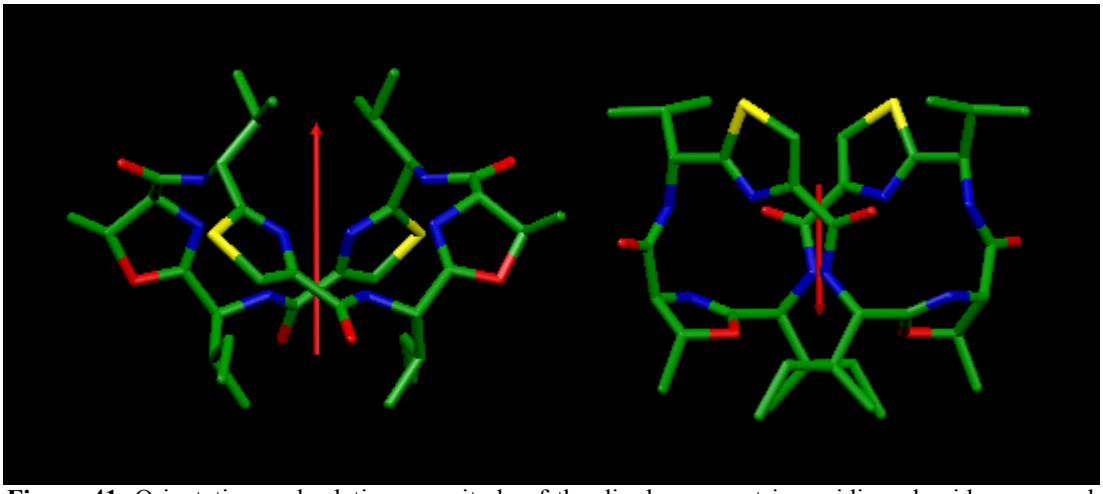


Figure 41. Orientation and relative magnitude of the dipole moment in ascidiacyclamide open and closed conformations. Hydrogens omitted to improve clarity.

Coordination geometries and reactivities of patellamide TAO/Cu^{2+} complexes

The studies of the coordination geometries of the TAO/Cu^{2+} complexes (Chapter 4) showed that with the neutral ligands H_2O and MeOH the preferred primary coordination number for Cu^{2+} in this environment is five. MeOH was included in the study as this had been the solvent used in previous experimental work on these complexes. The equilibrium geometry for complexes with these ligands is found to be square pyramidal/trigonal bipyramidal distorted away from the central cavity of the peptide macrocycle. On moving to the uninegative ligands Cl^- and OH^- a change in the primary coordination sphere of the copper is observed with a strong directing effect forcing the negatively changed ligands to take up a position trans to the deprotonated amide nitrogen in the plane of the complex (in the central macrocyclic cavity in the full patellamide structure). The equilibrium geometry of the chloride-containing complex is square pyramidal whilst that of the hydroxide-containing complex is square planar. From consideration of the effects on the coordination environment due to these ligands it seems possible that the coordinative desaturation of the copper in the latter case may lead to enhanced reactivity as the metal is much more exposed to attack from incoming reactants.

Through the use of the DFT local reactivity indices provided by the Fukui function,[109] $f(\vec{r})$, the reactivity of the copper in the TAO site with the exogenous ligands mentioned above was investigated (Chapter 5). The differential form of $f(\vec{r})$ is shown below

$$f(\vec{r}) = \left(\frac{\partial \rho(\vec{r})}{\partial N} \right)_{v(\vec{r})}$$

The reactivity of the copper in the H_2O- and MeOH-containing complexes was found to be almost identical (Table 31). This, along with the similar coordination geometry found for these complexes goes some way to support the continued use of MeOH in attempts to experimentally determine the in-vitro function of the patellamide/copper systems.

		Global Properties			Local Properties		
Complex	q_{Cu}^{M}	I	A	η	f_{Cu}^{+}	f_{Cu}^{-}	f_{Cu}^{0}
$[Cu(H_2O)_5]^{2+}$	1.181	19.952	9.050	5.451	0.532	0.172	0.352
$[Cu(phen)(H_2O)_3]^{2+}$	1.127	14.708	8.033	3.338	0.374	0.047	0.210
$[Cu(TAO)(H_2O)_2]^{+}$	1.026	10.678	3.676	3.501	0.356	0.101	0.228
$[Cu(TAO)(MeOH)_2]^{+}$	1.054	10.497	3.724	3.386	0.341	0.092	0.216
[Cu(TAO)(H_2O)Cl]	0.932	7.523	0.048	3.738	0.264	0.061	0.162
[Cu(TAO)(OH)]	0.930	7.117	0.381	3.368	0.360	0.151	0.255

Table 31. Summary of values for reactivity descriptors calculated at B3LYP/6-31G(d,p) level. Mulliken partial charge (q_{Cu}^{M}), ionization potential (I, eV), electron affinity (A, eV), global hardness (η) and condensed Fukui functions (f_{Cu}).

The results of the calculations for the chloride and hydroxide systems show large differences in the reactivity of the copper center, with chloride acting as an inhibitor and hydroxide enhancing the reactivity. This is in agreement with the observations of experimentalists working on the possibility of a catalytic rôle for the patellamide/copper systems who have noted a reduction in activity in the presence of dissolved Cl^- and an increase in activity in basic solutions (where the concentration of hydroxide would be expected to be significant). The high value of f_{Cu} for nucleophilic susceptibility in the complex containing hydroxide ligand may point to some degree of selectivity for the nature of the nucleophile as during the geometry optimizations the aquo ligands were ejected despite being (weakly) nucleophilic.

Because of the use of a reduced model of the patellamide structure in these studies it is felt

that future studies in this area should concentrate on elucidating the nature of the copper binding in the full patellamide structure, as steric hinderance by the amino acid side-chains and features of the macrocyclic cavity are likely to be important in mediating the binding of exogenous ligands. A further reason for including the full peptide is that this will allow the investigation of the environment of the dicopper center observed in the crystal structure of ASC.[69]

References

(1) Young, J. Z. *The Life of Vertebrates*; 2nd ed.; Oxford University Press: London, 1962.
(2) Borradaile, L. A.; Eastham, L. E. S.; Potts, F. A.; Saunders, J. T. *The Invertebrata*; 4th ed.; Cambridge University Press: London, 1963.
(3) Talbot, F. (scientific consultant). *The Reader's Digest Book of the Great Barrier Reef*; 1st ed.; Reader's Digest: Sydney, 1984.
(4) Salomon, C. E.; Faulkner, D. J. *J. Nat. Prod.* **2002**, *65*, 689-692.
(5) Degnan, B. M.; Hawkins, C. J.; Lavin, M. F.; McCaffrey, E. J.; Parry, D. L.; Vandenbrenk, A. L.; Watters, D. J. *J. Med. Chem.* **1989**, *32*, 1349-1354.
(6) Degnan, B. M.; Hawkins, C. J.; Lavin, M. F.; McCaffrey, E. J.; Parry, D. L.; Watters, D. J. *J. Med. Chem.* **1989**, *32*, 1354-1359.
(7) Schmidt, U.; Weller, D. *Tet. Lett.* **1986**, *27*, 3495-3496.
(8) Kato, S.; Hamada, Y.; Shioiri, T. *Tet. Lett.* **1986**, *27*, 2653-2656.
(9) Williams, D. E.; Moore, R. E.; Paul, V. J. *J. Nat. Prod.* **1989**, *52*, 732-739.
(10) Wasylyk, J. M.; Biskupiak, J. E.; Costello, C. E.; Ireland, C. M. *J. Org. Chem.* **1983**, *48*, 4445-4449.
(11) Morris, L. A.; van den Bosch, J. J. K.; Versluis, K.; Thompson, G. S.; Jaspars, M. *Tetrahedron* **2000**, *56*, 8345-8353.
(12) Roy, R. S.; Gehring, A. M.; Milne, J. C.; Belshaw, P. J.; Walsh, C. T. *Nat. Prod. Rep.* **1999**, *16*, 249-263.
(13) Ishida, T.; Tanaka, M.; Nabae, M.; Inoue, M.; Kato, S.; Hamada, Y.; Shioiri, T. *J. Org. Chem.* **1988**, *53*, 107-112.
(14) In, Y.; Doi, M.; Inoue, M.; Ishida, T.; Hamada, Y.; Shioiri, T. *Chem. Pharm. Bull.* **1993**, *41*, 1686-1690.
(15) Ishida, T.; In, Y.; Shinozaki, F.; Doi, M.; Yamamoto, D.; Hamada, Y.; Shioiri, T.; Kamigauchi, M.; Sugiura, M. *J. Org. Chem.* **1995**, *60*, 3944-3952.
(16) Doi, M.; Shinozaki, F.; In, Y.; Ishida, T.; Yamamoto, D.; Kamigauchi, M.; Sugiura, M.; Hamada, Y.; Kohda, K.; Shioiri, T. *Biopolymers* **1999**, *49*, 459-469.
(17) Milne, B. F.; Morris, L. A.; Jaspars, M.; Thompson, G. S. *J. Chem. Soc.-Perkin Trans. 2* **2002**, 1076-1080.
(18) Grondahl, L.; Sokolenko, N.; Abbenante, G.; Fairlie, D. A.; Hanson, G. R.; Gahan, L. R. *J. Chem. Soc.-Dalton Trans.* **1999**, 1227-1234.
(19) Freeman, D. J.; Pattenden, G.; Drake, A. F.; Siligardi, G. *J. Chem. Soc.-Perkin Trans. 2* **1998**, 129-135.
(20) Morris, L. A.; Milne, B. F.; Thompson, G. S.; Jaspars, M. *J. Chem. Soc.-Perkin Trans. 2* **2002**, 1072-1075.
(21) Ireland, C. M.; Durso, A. R.; Newman, R. A.; Hacker, M. P. *J. Org. Chem.* **1982**, *47*, 1807-1811.
(22) Williams, A. B.; Jacobs, R. S. *Cancer Lett.* **1993**, *71*, 97-102.
(23) Fu, X.; Do, T.; Schmitz, F. J.; Andrusevich, V.; Engel, M. H. *J. Nat. Prod.* **1998**, *61*, 1547-1551
(24) Michael, J. P.; Pattenden, G. *Angew. Chem. Int. Ed. (English)* **1993**, *32*, 1-23.
(25) Schmitz, F. J.; Ksebati, M. B.; Chang, J. S.; Wang, J. L.; Hossain, M. B.; Vanderhelm, D.; Engel, M. H.; Serban, A.; Silfer, J. A. *J. Org. Chem.* **1989**, *54*, 3463-3472.
(26) Ishida, T.; Inoue, M.; Hamada, Y.; Kato, S.; Shioiri, T. *J. Chem. Soc.-Chem. Commun.* **1987**, 370-371.

(27) Ishida, T.; In, Y.; Doi, M.; Inoue, M.; Hamada, Y.; Shioiri, T. *Biopolymers* **1992**, *32*, 131-143.
(28) In, Y.; Doi, M.; Inoue, M.; Ishida, T.; Hamada, Y.; Shioiri, T. *Acta Crystallogr. Sect. C-Cryst. Struct. Commun.* **1994**, *50*, 2015-2017.
(29) Wipf, P.; Fritch, P. C.; Geib, S. J.; Sefler, A. M. *J. Am. Chem. Soc.* **1998**, *120*, 4105-4112.
(30) GAMESS-UK is a package of ab initio programs written by M.F. Guest, J.H. van Lenthe, J. Kendrick, K. Schoffel, and P. Sherwood, with contributions from R.D. Amos, R.J. Buenker, H.J.J. van Dam, M. Dupuis, N.C. Handy, I.H. Hillier, P.J. Knowles, V. Bonacic-Koutecky, W. von Niessen, R.J. Harrison, A.P. Rendell, V.R. Saunders, A.J. Stone and A.H. de Vries. The package is derived from the original GAMESS code due to M. Dupuis, D. Spangler and J. Wendoloski, NRCC Software Catalog, Vol. 1, Program No. QG01 (GAMESS), 1980.
(31) Bone, S. F.; Smart, B. A.; Geirens, H.; Morrison, C. A.; Brain, P. T.; Rankin, D. W. H. *Phys. Chem. Chem. Phys.* **1999**, *1*, 2421-2426.
(32) Nygaard, L.; Asmussen, E.; Hog, J. H.; Maheshwari, R. C.; Nielsen, C. H.; Petersen, I. B.; Rastrup-Andersen, J.; Sorensen, G. *J. Mol. Struc.* **1971**, *8*, 225-233.
(33) El-Azhary, A. A. *J. Chem. Res.-S* **1995**, 174-175.
(34) El-Azhary, A. A.; Ghoneim, A. A.; Elshakre, M. E. *J. Chem. Res.-S* **1995**, 354-355.
(35) Kassimi, N. E.; Doerksen, R. J.; Thakkar, A. J. *J. Phys. Chem.* **1996**, *100*, 8752-8757.
(36) Kassimi, N. E. B.; Lin, Z. J. *J. Phys. Chem. A* **1998**, *102*, 9906-9911.
(37) El-Azhary, A. A.; Suter, H. U. *J. Phys. Chem.* **1996**, *100*, 15056-15063.
(38) El-Azhary, A. A. *J. Chem. Res.-M* **1995**, 1146-1154.
(39) Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*; 1st ed.; Wiley-VCH: Weinheim, 2000.
(40) Jensen, F. *Introduction to Computational Chemistry*; 1st ed.; John Wiley and Sons Limited: Chichester, 1999.
(41) Stephens, P. J.; Devlin, F. J.; Chablowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623-11627.
(42) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.
(43) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098-3100.
(44) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.
(45) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785-789.
(46) Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. *J. Chem. Phys.* **1998**, *109*, 6264-6271.
(47) Altmann, J. A.; Handy, N. C. *Phys. Chem. Chem. Phys.* **1999**, *1*, 5529-5536.
(48) Stevens, W. J.; Krauss, M.; Basch, H.; Jasien, P. G. *Can. J. Chem.* **1992**, *70*, 612-630.
(49) Windholz, M.; Budavari, S.; Stroumtsos, L. Y.; Fertig, M. N., Eds. *The Merck Index*; 9th ed.; Merck & Co. Inc.: Rahway, N.J., U.S.A., 1976.
(50) Grev, R. S.; Schaefer (III), H. F. *J. Chem. Phys.* **1989**, *91*, 7305-7306.
(51) Johnson, B. G.; Gill, P. M. W.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 5612.
(52) Bauschlicher, C. W.; Partridge, H. *J. Chem. Phys.* **1995**, *103*, 1788-1791.
(53) Montgomery Jr., J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1999**, *110*, 2822-2827.
(54) Jishi, R. A.; Flores, R. M.; Valderrama, M.; Lou, L.; Bragin, J. *J. Phys. Chem. A* **1998**, *102*, 9858-9862.

(55) Patel, M. A.; Derety, E.; Csizmadia, I. G. *J. Mol. Struc. (Theochem)* **1999**, *492*, 1-18.
(56) Rauk, A.; Armstrong, D. A. *J. Am. Chem. Soc.* **2000**, *122*, 4185-4192.
(57) Gunther, R.; Hofmann, H. J. *J. Am. Chem. Soc.* **2001**, *123*, 247-255.
(58) Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. *Chem. Phys.* **2001**, *263*, 203-219.
(59) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caulfield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.
(60) Slifkin, M. A. In *Molecular Interactions*; 1st ed.; Ratajczak, H., Orville-Thomas, W. J., Eds.; John Wiley & Sons: Chichester, 1981; Vol. 2, pp 271-304.
(61) Parr, R. G.; Pearson, R. G. *J. Am. Chem. Soc.* **1983**, *105*, 7512-7516.
(62) Pearson, R. G. *J. Am. Chem. Soc.* **1985**, *107*, 6801-6806.
(63) Chattaraj, P. K.; Parr, R. G. In *Structure and Bonding*; Springer-Verlag: Heidelberg, 1993; Vol. 80 (Chemical Hardness), pp 11-26.
(64) Parr, R. G.; Donnelly, R. A.; Levy, M.; Palke, W. E. *J. Chem. Phys.* **1978**, *68*, 3801-3807.
(65) Berkowitz, M.; Ghosh, S. K.; Parr, R. G. *Ann. Rev. Phys. Chem.* **1985**, *34*, 6811-6814.
(66) Pearson, R., G. In *Structure and Bonding*; Springer-Verlag: Heidelberg, 1993; Vol. 80 (Chemical Hardness), pp 1-10.
(67) Parr, R. G.; Sventpaly, L. v.; Liu, S. *J. Am. Chem. Soc.* **1999**, *121*, 1922-1924.
(68) Morris, L. A. *Studies on the Cu(II) and Zn(II) Binding Properties of Cyclic Peptides from the Ascidian Lissoclinum patella*; PhD. Thesis, Department of Chemistry, University of Aberdeen: Aberdeen, 1999.
(69) van den Brenk, A. L.; Byriel, K. A.; Fairlie, D. A.; Gahan, L. R.; Hanson, G. R.; Hawkins, C. J.; Jones, A.; Kennard, C. H. L.; Moubaraki, B.; Murray, K. S. *Inorg. Chem.* **1994**, *33*, 3549-3557.
(70) van den Brenk, A. L.; Fairlie, D. P.; Hanson, G. R.; Gahan, L. R.; Hawkins, C. J.; Jones, A. *Inorg. Chem.* **1994**, *33*, 2280-2289.
(71) Morris, L. A.; Jaspars, M.; Kettenes-van den Bosch, J. J.; Versluis, K.; Heck, A. J. R.; Kelly, S. M.; Price, N. C. *Tetrahedron* **2001**, *57*, 3185-3197.
(72) Cusack, R. M.; Grondahl, L.; Abbenante, G.; Fairlie, D. P.; Gahan, L. R.; Hanson, G. R.; Hambley, T. W. *J. Chem. Soc.-Perkin Trans. 2* **2000**, 323-331.
(73) Bernhardt, P. V.; Comba, P.; Fairlie, D. P.; Gahan, L. R.; Hanson, G. R.; Lotzbeyer, L. *Chem.-Eur. J.* **2002**, *8*, 1527-1536.
(74) Housecroft, C. E.; Sharpe, A. G. *Inorganic Chemistry*; 1st ed.; Pearson Education Limited: Harlow, 2001.
(75) Greenwood, N. N.; Earnshaw, A. *Chemistry of the Elements*; 8th ed.; Pergamon Press Ltd.: Oxford, 1984.
(76) Lippard, S. J.; Berg, J. M. *Principles of Bioinorganic Chemistry*; 1st ed.; University Science Books: Mill Valley, CA 94941, 1994.
(77) Kaim, W.; Schwederski, B. *Bioinorganic Chemistry: Inorganic Elements in the Chemistry of Life*; 6th ed.; John Wiley & Sons Ltd.: Chichester, 1991.
(78) Abbenante, G.; Fairlie, D. P.; Gahan, L. R.; Hanson, G. R.; Pierens, G. K.; van den Brenk, A. L. *J. Am. Chem. Soc.* **1996**, *118*, 10384-10388.
(79) Burgess, J. *Metal Ions in Solution*; 1st ed.; John Wiley & Sons Ltd.: Chichester, 1978.

(80) Solomon, E. I.; Sundaram, U. M.; Machonkin, T. E. *Chem. Rev.* **1996**, *96*, 2563-2605.
(81) Decker, H.; Terwilliger, N. *J. Exp. Biol.* **2000**, *203*, 1777-1782.
(82) Mayer, F.; Ackerman, J.; Pritzkow, H. In *Proceedings of the 34th International Conference on Coordination Chemistry*; Edinburgh, Scotland, July 2000.
(83) Glaser, O.; Anslow, G. A. *J. Exp. Zool.* **1949**, *111*, 117-139.
(84) Bellas, J.; Vazquez, E.; Beiras, R. *Water Res.* **2001**, *35*, 2905-2912.
(85) Raftos, D.; Hutchinson, A. *Biol. Bull.* **1997**, *192*, 62-72.
(86) Arizza, V.; Cammarata, M.; Tomasino, M. C.; Parrinello, N. *J. Invertebr. Pathol.* **1995**, *66*, 297-302.
(87) Radford, J. L.; Hutchinson, A. E.; Burandt, M.; Raftos, D. A. *J. Invertebr. Pathol.* **2000**, *76*, 242-248.
(88) Tujula, N.; Radford, J.; Nair, S. V.; Raftos, D. A. *Aquat. Toxicol.* **2001**, *55*, 191-201.
(89) Hata, S.; Azumi, K.; Yokosawa, H. *Comp. Biochem. Physiol. B-Biochem. Mol. Biol.* **1998**, *119*, 769-776.
(90) Cima, F.; Perin, A.; Burighel, P.; Ballarin, L. *Acta Zool.* **2001**, *82*, 261-274.
(91) Comba, P.; Cusack, R.; Fairlie, D. P.; Gahan, L. R.; Hanson, G. R.; Kazmaier, U.; Ramlow, A. *Inorg. Chem.* **1998**, *37*, 6721-6727.
(92) Bernhardt, P. V.; Comba, P. *Inorg. Chem.* **1992**, *31*, 2638-2644.
(93) Bauschlicher, C. W. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Kollman, P. A., Clark, T., Schaefer III, H. F., Gasteiger, J., Schreiner, P. R., Eds.; John Wiley & Sons Ltd.: Chichester, 1998; pp 3084-3094.
(94) In, Y.; Noi, M.; Inoue, M.; Ishida, T. *Acta Crystallogr. Sect. C-Cryst. Struct. Commun.* **1994**, *50*, 432-434.
(95) Ryde, U.; Olsson, M. H. M.; Pierloot, K.; Roos, B. O. *J. Mol. Biol.* **1996**, *261*, 586-596.
(96) Haeffner, F.; Brink, T.; Haeberlein, M.; Moberg, C. *J. Mol. Struc. (Theochem)* **1997**, *397*, 39-50.
(97) Cotton, F. A.; Feng, X.; Timmons, D. J. *Inorg. Chem.* **1998**, *37*, 4066-4069.
(98) Sabolovic, J.; Liedl, K. R. *Inorg. Chem.* **1999**, *38*, 2764-2774.
(99) Flock, M.; Pierloot, K. *J. Phys. Chem. A* **1999**, *103*, 95-102.
(100) Mastryukov, V.; Ruan, C. Y.; Fink, M.; Wang, Z.; Pachter, R. *J. Mol. Struc.* **2000**, *556*, 225-237.
(101) Martinez, A.; Salcedo, R.; Sansores, L. E.; Medina, G.; Gasque, L. *Inorg. Chem.* **2001**, *40*, 301-306.
(102) Legge, F. S.; Nyberg, G. L.; Peel, J. B. *J. Phys. Chem. A* **2001**, *105*, 7905-7916.
(103) Fomina, L.; Vazquez, B.; Tkatchouk, E.; Fomine, S. *Tetrahedron* **2002**, *58*, 6741-6747.
(104) Morris, L. A.; Milne, B. F.; Jaspars, M.; Kettenes-van den Bosch, J. J.; Versluis, K.; Heck, A. J. R.; Kelly, S. M.; Price, N. C. *Tetrahedron* **2001**, *57*, 3199-3207.
(105) Muguruma, C.; Koga, N.; Kitaura, K.; Morokuma, K. *J. Chem. Phys.* **1995**, *103*, 9274-9291.
(106) Dediu, A. *Chem. Rev.* **2000**, *100*, 543-600.
(107) Sigel, H.; Huber, P. H.; Griesser, R.; Prijs, B. *Inorg. Chem.* **1973**, *12*, 1198-1200.
(108) Yang, W.; Parr, R. G. *Proc. Nat. Acad. Sci. (USA)* **1985**, *82*, 6723-6726.
(109) Parr, R. G.; Yang, W. *J. Am. Chem. Soc.* **1984**, *106*, 4049-4050.
(110) Yang, W.; Mortier, W. J. *J. Am. Chem. Soc.* **1986**, *108*, 5708-5711.

Appendix

Publications

Previously Published Work

All molecular modelling work in the following publications was carried out by the present author during the allotted Ph.D. study period (with the exception of the nOe-restrained MD studies of the metal bound conformational properties of the patellamides contained in Morris *et al.*, *J. Chem. Soc., Perkin Trans. 2*, 2002, 1072-1075 which were performed by Dr. Marcel Jaspars (University of Aberdeen, UK) and Dr. Gary S. Thompson (University of Leeds, UK)).

Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Crystallographic and molecular mechanics investigation of an order–disorder transition and dimorphism in 5*H*,10*H*-dithiolo[2,3-*b*]-2,5-benzodithiocine-2-thione

Zahid H. Chohan, William T. A. Harrison, R. Alan Howie, Bruce F. Milne and James L. Wardell*

Department of Chemistry, University of Aberdeen, Old Aberdeen AB24 3UE, Scotland

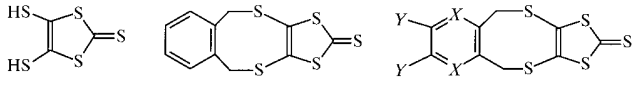
Correspondence e-mail: j.wardell@abdn.ac.uk

Received 24 February 2000
Accepted 28 July 2000

Single-crystal X-ray structures are presented for three forms of 5*H*,10*H*-dithiolo[2,3-*b*]-2,5-benzodithiocine-2-thione. The α (at 150 K) and α' (at ambient) forms are very similar and differ only in the presence of crystallographic *m* symmetry in the molecules of α', which is absent in the case of α. This pair is related by an order–disorder transition. The β phase (also determined at 150 K) has a different structure in terms of the molecular packing from either of the other two and therefore constitutes a true polymorph. Molecular mechanics calculations indicated that the most stable $CHCl_3$-solvated conformations for the title compound were a pair of twisted *U*-shaped enantiomers, U_R and U_L, *i.e.* similar to the arrangements found in the α and β phases, with the low-lying saddle point between them corresponding to the situation in the α' phase. These calculations also indicated that the most stable $CHCl_3$-solvated conformation for the related dibromo-5*H*,10*H*-dithiolo[2,3-*b*]-2,5-benzodithiocine-2-thione was *Z*-shaped, in agreement with the crystal structure determined earlier for its DMSO solvate [Wang *et al.* (1998). *Synthesis*, pp. 1615–1618].

1. Introduction

The packing of molecules in the crystalline state has a crucial bearing on their material properties. Various types of intermolecular and interion interactions can help fashion the solid-state assembly (Desiraju, 1995). In aromatic and other π-delocalized systems, π–π interactions are particularly important intermolecular forces (Hunter, 1994). For derivatives of 4,5-dimercapto-1,3-dithiole-2-thione [H_2-dmit: (1)], the frequently reported non-covalent intermolecular and interionic interactions are S···S contacts, as illustrated by the dmit compounds listed in the Cambridge Structural Database (Allen & Kennard, 1993; Fletcher *et al.*, 1995). As part of an ongoing study on dmit compounds, the structure of 5*H*,10*H*-dithiole[2,3-*b*]-2,5-benzodithiocine-2-thione (2) was determined at ambient temperature; surprisingly, this was found to have a *U*-shaped molecular structure in contrast to the *Z*-shaped structures established for the solvates of related compounds (3) [(3):DMSO] and (4) [(4):DMF] (Wang *et al.*, 1998). This finding led us onto a molecular mechanics study of (2) and (3), and the determination of structures at 150 K of (2)



(1) (2) (3): X = CH; Y = Br (4): X = N; Y = CN

recrystallized from different media. As we now report, the study revealed an order–disorder relationship and dimorphism in (2).

Table 1
Experimental details.

	α-(2)	α′-(2)	β-(2)
Crystal data			
Chemical formula	$C_{11}H_8S_5$	$C_{11}H_8S_5$	$C_{11}H_8S_5$
Chemical formula weight	300.47	300.47	300.47
Cell setting	Monoclinic	Monoclinic	Monoclinic
Space group	$P2_1/a$	$C2/m$	$P2_1/n$
a (Å)	11.9290 (2)	12.075 (8)	8.5683 (2)
b (Å)	11.7639 (2)	11.825 (8)	16.2080 (4)
c (Å)	9.0612 (2)	9.063 (6)	17.8980 (4)
β (°)	103.9498 (10)	102.80 (5)	97.1103 (14)
V (Å^3)	1234.07 (4)	1261.9 (15)	2466.47 (10)
Z	4	4	8
D_x (Mg m^{-3})	1.617	1.582	1.618
Radiation type	Mo $K\alpha$	Mo $K\alpha$	Mo $K\alpha$
Wavelength (Å)	0.71073	0.71073	0.71073
No. of reflections for cell parameters	21 934	14	31 494
θ range (°)	2.32–27.46	10.8–12.9	1.70–26.43
μ (mm^{-1})	0.905	0.885	0.905
Temperature (K)	150 (2)	298 (2)	150 (2)
Crystal form	Block	Block	Block
Crystal size (mm)	0.30 × 0.20 × 0.20	0.60 × 0.46 × 0.40	0.20 × 0.10 × 0.10
Crystal colour	Red	Pale brown	Orange
Data collection			
Diffractometer	Enraf–Nonius KappaCCD area detector	Nicolet P3	Enraf–Nonius KappaCCD area detector
Data collection method	ω scans	θ–2θ scans	ω scans
Absorption correction	Multi-scan	ψ	Multi-scan
T_{min}	0.728	0.619	0.809
T_{max}	0.835	0.719	0.914
No. of measured reflections	21 934	2034	31 494
No. of independent reflections	2825	1932	5044
No. of observed reflections	2500	1166	3825
Criterion for observed reflections	$I > 2\sigma(I)$	$I > 2\sigma(I)$	$I > 2\sigma(I)$
R_{int}	0.0472	0.0238	0.0734
θ_{max} (°)	27.46	30.09	26.43
Range of h, k, l	$-15 \rightarrow h \rightarrow 15$ $-15 \rightarrow k \rightarrow 15$ $-11 \rightarrow l \rightarrow 11$	$-16 \rightarrow h \rightarrow 16$ $-16 \rightarrow k \rightarrow 0$ $0 \rightarrow l \rightarrow 12$	$-10 \rightarrow h \rightarrow 10$ $-20 \rightarrow k \rightarrow 20$ $-22 \rightarrow l \rightarrow 22$
No. of standard reflections	–	2	–
Frequency of standard reflections	–	Every 50 reflections	–
Refinement			
Refinement on	F^2	F^2	F^2
$R[F^2>2\sigma(F^2)]$	0.0278	0.0862	0.0363
$wR(F^2)$	0.0712	0.2153	0.0918
S	1.051	1.053	1.032
No. of reflections used in refinement	2825	1932	5044
No. of parameters used	145	78	289
H-atom treatment	H-atom parameters constrained	Only H-atom U's refined	H-atom parameters constrained
Weighting scheme	$w = 1/[\sigma^2(F_o^2) + (0.0370P)^2 + 0.4749P]$, where $P = (F_o^2 + 2F_c^2)/3$	$w = 1/[\sigma^2(F_o^2) + (0.0535P)^2 + 8.1837P]$, where $P = (F_o^2 + 2F_c^2)/3$	$w = 1/[\sigma^2(F_o^2) + (0.0422P)^2 + 0.3638P]$, where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max}$	0.001	0.000	0.001
$\Delta\rho_{max}$ (e Å^{-3})	0.38	0.92	0.45
$\Delta\rho_{min}$ (e Å^{-3})	−0.38	−0.95	−0.29
Extinction method	None	None	None
Source of atomic scattering factors	*International Tables for Crystallography* (1992, Vol. C, Tables 4.2.6.8 and 6.1.1.4)	*International Tables for Crystallography* (1992, Vol. C, Tables 4.2.6.8 and 6.1.1.4)	*International Tables for Crystallography* (1992, Vol. C, Tables 4.2.6.8 and 6.1.1.4)
Computer programs			
Data collection	*DENZO* (Otwinowski & Minor, 1997), *COLLECT* (Hooft, 1998)	*P*3 software (Nicolet, 1980)	*DENZO* (Otwinowski & Minor, 1997), *COLLECT* (Hooft, 1998)
Cell refinement	*DENZO* (Otwinowski & Minor, 1997), *COLLECT* (Hooft, 1998)	*P*3 software (Nicolet, 1980)	*DENZO* (Otwinowski & Minor, 1997), *COLLECT* (Hooft, 1998)

Table 1 (continued)

	α-(2)	α'-(2)	β-(2)
Data reduction	*DENZO* (Otwinowski & Minor, 1997), *COLLECT* (Hooft, 1998)	*RDNIC* (Howie, 1980)	*DENZO* (Otwinowski & Minor, 1997), *COLLECT* (Hooft, 1998)
Structure solution	*SHELXS*97 (Sheldrick, 1990)	*SHELXS*97 (Sheldrick, 1990)	*SHELXS*97 (Sheldrick, 1990)
Structure refinement	*SHELXL*97 (Sheldrick, 1997)	*SHELXS*97 (Sheldrick, 1990)	*SHELXS*97 (Sheldrick, 1990)
Preparation of material for publication	*SHELXS*97 (Sheldrick, 1990)	*SHELXS*97 (Sheldrick, 1990)	*SHELXS*97 (Sheldrick, 1990)

2. Experimental

2.1. General techniques

Melting points were measured on a Kofler hot-stage microscope and are uncorrected. NMR spectra were recorded on Bruker 250 MHz and 400 MHz instruments, UV–vis spectra on a Perkin Elmer Lambda 15 instrument and IR spectra on a Philips Analytical PU9800 FTIR instrument. Differential scanning calorimetry was carried out on a DSC Metler Toledo 820 instrument. The compound $[NEt_4]_2[Zn(dmit)_2]$ was obtained by a published procedure (Wang *et al.*, 1998).

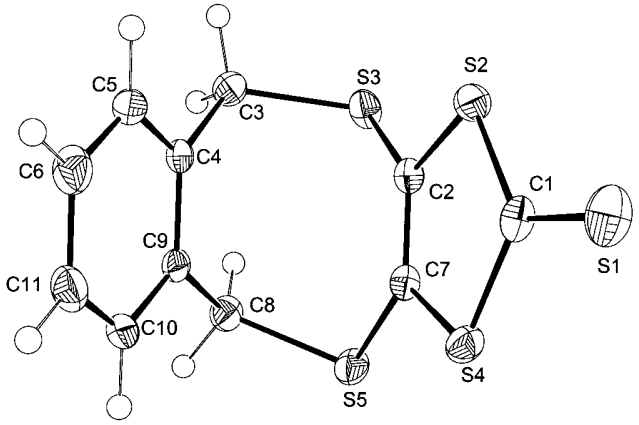


Figure 1
Atom-labelling scheme for α-(2) (50% probability displacement ellipsoids). H atoms represented by plain spheres of arbitrary radius.

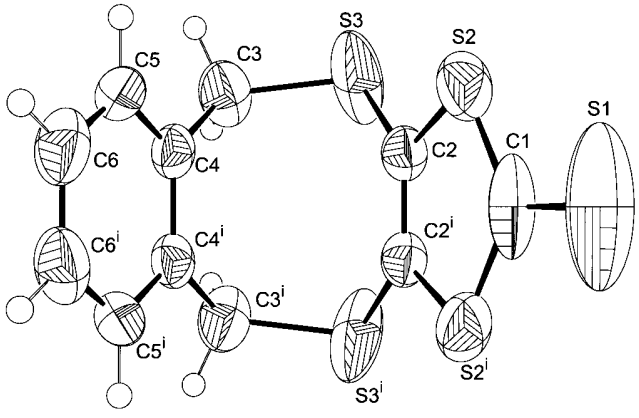


Figure 2
Atom-labelling scheme for α'-(2) (50% probability displacement ellipsoids). H atoms represented by plain spheres of arbitrary radius. Symmetry code: (i) $x, -y, z$.

2.2. Synthesis

Solutions of 1,2-bis(chloromethyl)benzene (0.88 g, 5.0 mmol) in CH_2Cl_2 (20 ml) and $[NEt_4]_2[Zn(dmit)_2]$ (1.80 g, 2.50 mmol) in CH_2Cl_2 (30 ml) were mixed and refluxed overnight. The solvent was removed to leave an oil, which was chromatographed on silica using CH_2Cl_2 as an eluent. The title compound was recrystallized from CH_2Cl_2/hexane as a red crystalline solid (polymorph α), m.p. 479 K [literature values 478 K (Kumar *et al.*, 1991), 480 K (Goldenberg & Lyubovskaya, 1986)]; yield 0.78 g, 52%. Recrystallization from DMF and $CHCl_3$ also gave polymorph α. Analysis: found: C 43.1, H 2.5%; calculated for $C_{11}H_8S_5$: C 43.2, H 2.7%. IR (KBr): ν(C═S) 1066 cm^{-1}. UV–vis (CH_2Cl_2): 374, 520 nm. ^{13}C NMR ($CDCl_3$, 63 MHz): δ 38.6 [CH_2], 129.0 [C4], 130.7 [C3], 134.0 [C1], 139.3 [C═C], 211.8 [C═S].

Crystallization of the residue of a reaction mixture also containing $Cr(CO)_6$ from ethanol gave the orange β phase, m.p. 479–480 K, whose solution NMR and solid-state IR spectral data were identical with those of the α' phase.

2.3. Molecular mechanics calculations

The calculation of the minimum energy solution ($CHCl_3$) conformations of (2) and (3) was carried out using Macromodel v6.5 (Mohamadi *et al.*, 1990) on a Silicon Graphics O2 workstation. Monte Carlo searching of the conformational spaces accessible to the two molecules was performed with subsequent Polak–Ribere conjugate gradient energy minimization of the generated structures. The MM2* force-field was employed for all energy calculations along with the GB/SA solvent model (Still *et al.*, 1990) using non-bonded cutoffs of 12 Å for electrostatic interactions and 7 Å for van der Waals' interactions.

2.4. Crystallography

Intensity data for α-(2) and β-(2) were collected at 150 K on an Enraf–Nonius KappaCCD area detector diffractometer, as outlined in Table 1.[1] Room-temperature data for α'-(2) were collected on a Nicolet P3 diffractometer (Table 1). For α-(2) and β-(2), absorption corrections were made on the basis of multiply measured and symmetry-equivalent reflections using *SORTAV* (Blessing, 1997), while ψ scans were applied to the data for α'-(2).

[1] Supplementary data for this paper are available from the IUCr electronic archives (Reference: BM0027). Services for accessing these data are described at the back of the journal.

Each structure was solved by direct methods with *SHELXS*86 (Sheldrick, 1990) and optimized by full-matrix least-squares refinement against F^2 by using *SHELXL*97 (Sheldrick, 1997). H-atom locations were calculated geometrically and thereafter allowed to ride on their respective C atoms. For the α and β phases, U_{iso} for each H was set to 1.2 × U_{eq} of its attached C. For the α' phase, separate common group U_{iso} values were refined for methylene and aryl H.

3. Results and discussion

3.1. General

Compound (2) was obtained from the reaction between $[NEt_4]_2[Zn(dmit)_2]$ and 1,2-$(ClCH_2)_2C_6H_4$ in CH_2Cl_2 solution. Recrystallization from CH_2Cl_2, DMF or CH_2Cl_2/EtOH gave the same product, designated the α' phase, from structure determinations at 298 (2) K, and the α phase, from structure determinations at 150 (2) K. In contrast to (3), (2) did not form a DMF solvate on recrystallization from DMF solution. DSC experiments with the α/α' material indicated a very weak (0.42 kJ mol^{-1}), but reproducible endothermic event at 249 K either on cooling from 330 K or heating from 170 K. This thermal behaviour is consistent with a subtle molecular rearrangement on going from the α to the α' form of (2) (see below). The β phase, determined at 150 K, was obtained from a reaction mixture of (2) and $Cr(CO)_6$. The role of the chromium hexacarbonyl is unknown at this stage.

The atom-labelling schemes of α-(2) and α'-(2) are shown in Figs. 1 and 2. The atom-labelling scheme of α-(2) is also used for β-(2) with the additional use of suffixes *a* and *b* to identify the two independent molecules (Figs. 3 and 4). Selected bond distances and angles are listed in Table 2. These molecules have '*U*'-shaped structures, with the benzene ring and the dmit moiety as the legs of the '*U*' and the methylene C atoms and the attached S atoms as its base. The *U*-shaped structures are significantly different from the *Z*-shaped structures of the compounds (3):DMSO and (4):DMF (Wang *et al.*, 1998).

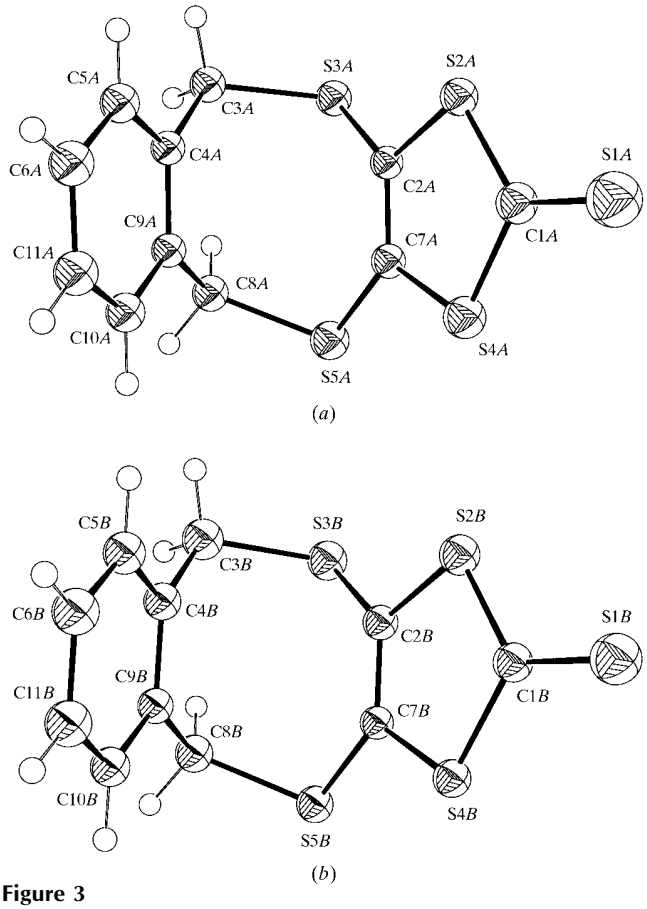


Figure 3
Atom-labelling scheme for (*a*) molecule *A* of β-(2) and (*b*) molecule *B* of β-(2) (50% probability displacement ellipsoids in each case). H atoms represented by plain spheres of arbitrary radius.

3.2. Molecular mechanics results

These calculations indicated five minima, as listed in Table 3. These minima corresponded to two enantiomeric and, therefore, equal-energied, twisted *U*-shaped conformations, U_L and U_R, a *Z*-shaped conformation and an enantiomeric pair of irregular conformations, T_L and T_R. The most stable $CHCl_3$-solvated conformation calculated for (3) is the *Z*-form, *i.e.* the conformation determined for the solid DMSO-solvate of (3). For molecule (2), the lowest energy conformations are, however, calculated to be the pair of enantiomers U_L and U_R, *i.e.* similar to the arrangements found in α-(2) and β-(2), with the low-lying saddle point between them corresponding to the situation in α'-(2), see below. The highest energy conformations, T_L and T_R, are twist forms in which the 'hinge' exocyclic S atoms lie one above and one below the phenyl ring: the angle between the planes of the phenyl ring and dmit ring is 61.6°.

The potential energy surface was investigated for the presence of saddle points lying between the minima using the *Macromodel* explicit saddle point search method (Culot *et al.*, 1990). Three saddle points (Table 4) were found to lie between true minima on the surface: (i) between U_L and U_R (a regular *U* shape) and (ii) two equal energy sites between U_L and T_L, and between U_R and T_R. Another saddle point was found to be

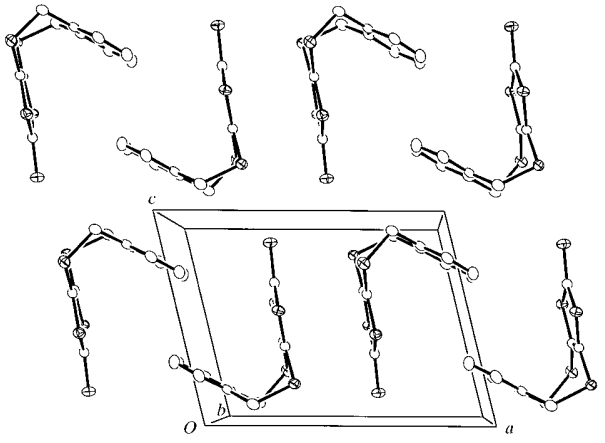


Figure 4
Packing-diagram for α-(2)

Table 2
Selected geometric parameters (Å, °).

α-2			
S1—C1	1.6400 (15)	C2—C7	1.354 (2)
S2—C1	1.7324 (16)	C3—C4	1.502 (2)
S2—C2	1.7449 (15)	C4—C5	1.393 (2)
S3—C2	1.7436 (14)	C4—C9	1.403 (2)
S3—C3	1.8477 (16)	C5—C6	1.383 (2)
S4—C1	1.7416 (17)	C6—C11	1.391 (3)
S4—C7	1.7442 (15)	C8—C9	1.498 (2)
S5—C7	1.7466 (15)	C9—C10	1.391 (2)
S5—C8	1.8468 (16)	C10—C11	1.383 (2)
C1—S2—C2	98.01 (7)	C9—C4—C3	122.86 (13)
C2—S3—C3	102.93 (7)	C6—C5—C4	121.47 (15)
C1—S4—C7	98.17 (7)	C5—C6—C11	119.57 (15)
C7—S5—C8	102.55 (7)	C2—C7—S4	115.53 (11)
S1—C1—S2	125.54 (10)	C2—C7—S5	128.61 (12)
S1—C1—S4	122.65 (10)	S4—C7—S5	115.85 (9)
S2—C1—S4	111.81 (8)	C9—C8—S5	111.76 (10)
C7—C2—S3	124.76 (12)	C10—C9—C4	119.20 (13)
C7—C2—S2	116.24 (11)	C10—C9—C8	119.10 (14)
S3—C2—S2	118.97 (9)	C4—C9—C8	121.67 (13)
C4—C3—S3	114.66 (10)	C11—C10—C9	121.43 (15)
C5—C4—C9	118.86 (14)	C10—C11—C6	119.46 (15)
C5—C4—C3	118.27 (14)		
α′-2			
S1—C1	1.635 (8)	C2—C2^{i}	1.334 (9)
S2—C1	1.719 (4)	C3—C4	1.506 (7)
S2—C2	1.724 (5)	C4—C5	1.376 (6)
S3—C2	1.735 (5)	C4—C4^{i}	1.392 (9)
S3—C3	1.773 (5)	C5—C6	1.380 (7)
C1—S2^{i}	1.719 (4)	C6—C6^{i}	1.384 (12)
C1—S2—C2	97.5 (3)	S2—C2—S3	117.2 (3)
C2—S3—C3	106.2 (2)	C4—C3—S3	115.9 (3)
S1—C1—S2^{i}	123.8 (2)	C5—C4—C4^{i}	119.4 (3)
S1—C1—S2	123.8 (2)	C5—C4—C3	118.6 (5)
S2^{i}—C1—S2	112.4 (4)	C4^{i}—C4—C3	122.1 (3)
C2^{i}—C2—S2	116.20 (16)	C4—C5—C6	121.2 (5)
C2^{i}—C2—S3	126.55 (18)	C5—C6—C6^{i}	119.5 (3)
β-2			
S1*A*—C1*A*	1.642 (3)	S1*B*—C1*B*	1.640 (2)
S2*A*—C1*A*	1.729 (3)	S2*B*—C1*B*	1.732 (3)
S2*A*—C2*A*	1.751 (2)	S2*B*—C2*B*	1.751 (2)
S3*A*—C2*A*	1.747 (2)	S3*B*—C2*B*	1.739 (2)
S3*A*—C3*A*	1.857 (2)	S3*B*—C3*B*	1.842 (3)
S4*A*—C1*A*	1.743 (3)	S4*B*—C1*B*	1.735 (3)
S4*A*—C7*A*	1.748 (2)	S4*B*—C7*B*	1.743 (2)
S5*A*—C7*A*	1.746 (3)	S5*B*—C7*B*	1.747 (2)
S5*A*—C8*A*	1.836 (2)	S5*B*—C8*B*	1.842 (3)
C2*A*—C7*A*	1.353 (3)	C2*B*—C7*B*	1.355 (3)
C3*A*—C4*A*	1.502 (3)	C3*B*—C4*B*	1.506 (3)
C4*A*—C5*A*	1.396 (3)	C4*B*—C5*B*	1.395 (3)
C4*A*—C9*A*	1.404 (3)	C4*B*—C9*B*	1.406 (3)
C5*A*—C6*A*	1.383 (4)	C5*B*—C6*B*	1.383 (4)
C6*A*—C11*A*	1.379 (4)	C6*B*—C11*B*	1.376 (4)
C8*A*—C9*A*	1.496 (3)	C8*B*—C9*B*	1.500 (3)
C9*A*—C10*A*	1.385 (3)	C9*B*—C10*B*	1.392 (3)
C10*A*—C11*A*	1.382 (4)	C10*B*—C11*B*	1.381 (3)
C1*A*—S2*A*—C2*A*	98.25 (12)	C1*B*—S2*B*—C2*B*	98.23 (12)
C1*A*—S2*A*—S4*B*	137.67 (9)	C2*B*—S3*B*—C3*B*	103.42 (11)
C2*A*—S2*A*—S4*B*	103.26 (8)	C1*B*—S4*B*—C7*B*	98.40 (12)
C2*A*—S3*A*—C3*A*	104.81 (11)	C1*B*—S4*B*—S2*A*	149.32 (9)
C1*A*—S4*A*—C7*A*	98.21 (12)	C7*B*—S4*B*—S2*A*	96.82 (8)
C7*A*—S5*A*—C8*A*	103.03 (11)	C7*B*—S5*B*—C8*B*	103.37 (11)
S1*A*—C1*A*—S2*A*	125.05 (17)	S1*B*—C1*B*—S2*B*	125.42 (15)
S1*A*—C1*A*—S4*A*	123.12 (16)	S1*B*—C1*B*—S4*B*	122.83 (16)
S2*A*—C1*A*—S4*A*	111.78 (14)	S2*B*—C1*B*—S4*B*	111.74 (13)
C7*A*—C2*A*—S3*A*	125.47 (19)	C7*B*—C2*B*—S3*B*	125.24 (19)
C7*A*—C2*A*—S2*A*	115.93 (18)	C7*B*—C2*B*—S2*B*	115.63 (18)
S3*A*—C2*A*—S2*A*	118.34 (14)	S3*B*—C2*B*—S2*B*	118.94 (14)
C4*A*—C3*A*—S3*A*	115.85 (17)	C4*B*—C3*B*—S3*B*	115.48 (18)
C5*A*—C4*A*—C9*A*	118.5 (2)	C5*B*—C4*B*—C9*B*	118.3 (2)
C5*A*—C4*A*—C3*A*	118.7 (2)	C5*B*—C4*B*—C3*B*	118.8 (2)
C9*A*—C4*A*—C3*A*	122.7 (2)	C9*B*—C4*B*—C3*B*	122.9 (2)
C6*A*—C5*A*—C4*A*	121.4 (2)	C6*B*—C5*B*—C4*B*	121.5 (2)
C11*A*—C6*A*—C5*A*	119.5 (2)	C11*B*—C6*B*—C5*B*	120.0 (2)
C2*A*—C7*A*—S5*A*	129.31 (19)	C2*B*—C7*B*—S4*B*	115.76 (18)
C2*A*—C7*A*—S4*A*	115.62 (19)	C2*B*—C7*B*—S5*B*	128.85 (19)
S5*A*—C7*A*—S4*A*	115.07 (14)	S4*B*—C7*B*—S5*B*	115.37 (14)
C9*A*—C8*A*—S5*A*	114.06 (17)	C9*B*—C8*B*—S5*B*	111.83 (17)
C10*A*—C9*A*—C4*A*	119.3 (2)	C10*B*—C9*B*—C4*B*	119.4 (2)
C10*A*—C9*A*—C8*A*	119.5 (2)	C10*B*—C9*B*—C8*B*	119.1 (2)
C4*A*—C9*A*—C8*A*	121.2 (2)	C4*B*—C9*B*—C8*B*	121.5 (2)
C11*A*—C10*A*—C9*A*	121.3 (2)	C11*B*—C10*B*—C9*B*	121.3 (2)
C6*A*—C11*A*—C10*A*	119.8 (2)	C6*B*—C11*B*—C10*B*	119.6 (2)

Symmetry codes: (i) $\frac{3}{2}-x, y-\frac{1}{2}, \frac{1}{2}-z$; (ii) $\frac{3}{2}-x, \frac{1}{2}+y, \frac{1}{2}-z$.

located between the *U* saddle point and the *Z* form. No direct route was found between *Z* and the T_L/T_R conformations. At the saddle point, $U_L \leftrightarrow U_R$ (= *U*), the thione group is somewhat bent towards the phenyl ring. The substitution of Br atoms onto the phenyl ring does not lead to a great change in the difference in energy between the conformations U_L/U_R and *Z* (< 1.25 kJ mol^{-1} in both molecules). Assuming a flat energy surface between the states, a relatively even distribution of conformers would be expected. However, the substitution increases the barrier to conformational inversion by ∼20 kJ mol^{-1} on going from U_L/U_R to *Z* and by ∼70 kJ mol^{-1} on going from *Z* to U_L/U_R, which is probably a steric effect. Conformations T_L and T_R are not affected much by this substitution and both the energy difference and the barrier to inversion remain within ∼2 kJ mol^{-1} of their original values.

3.3. Comparison of X-ray structures

The molecule in the *α*′-(2) polymorph has *m* symmetry, which is perpendicular to and bisects the benzene ring. Atoms S1 and C1 lie on this mirror plane (Fig. 2). The structure of *α*-(2) (Fig. 1), on the other hand, consists of two enantiomeric conformations (one corresponding to the asymmetric unit and one generated by space-group symmetry), in which S1 is displaced from the centroid, towards or away from C5, to give structures which are similar to the U_R and U_L forms calculated in the molecular mechanics study. The U_R- and U_L-type conformations are also observed for each of the two independent molecules of *β*-(2). As modelled here, the molecules of the asymmetric unit of *β*-(2) are both U_L types, but once again the structure is centrosymmetric. The calculated energy barrier for the U_R/U_L interconversion in (2) is only 6.55 kJ mol^{-1} and is therefore readily overcome at room temperature. Thus, if the temperature is high enough, the molecules can librate with interchange of conformation between the U_R and U_L states. It is for this reason that the disorder of structure *α*′-(2), relative to *α*-(2), is considered to be dynamic (librational) rather than static in nature.

In crystallographic terms, the relationship between structures *α*′-(2) and *α*-(2) is simple. The data collection and structure refinement of *α*-(2) was originally carried out in

Table 3
Energies and populations of conformational minima.

	(2)			(3)		
Minima	E_{MM2*} (kJ mol^{-1})	Boltzman factor	Population ratios (300 K)	E_{MM2*} (kJ mol^{-1})	Boltzman factor	Population ratios (300 K)
U_L	−55.22	4.28×10^9	389	−30.21	1.86×10^5	155
U_R	−55.22	4.28×10^9	389	−30.21	1.86×10^5	155
Z	−53.36	2.03×10^9	185	−31.10	2.66×10^5	222
T_L	−40.36	1.10×10^7	1	−17.64	1.20×10^3	1
T_R	−40.36	1.10×10^7	1	−17.64	1.20×10^3	1

Table 4
Energies of saddle points.

Saddle point	(2) E_{MM2*} (kJ mol^{-1})	(3) E_{MM2*} (kJ mol^{-1})
$[U_L \leftrightarrow U_R] = U$	−48.67	−23.06
$[U_L \leftrightarrow T_L]$	−29.15	−6.03
$[U_R \leftrightarrow T_R]$	−29.15	−6.03
$[U \leftrightarrow Z]$	−1.15	43.00

space group $P2_1/c$: subsequent re-refinement, after application of the appropriate cell and Miller index transformations, was carried out in the alternative non-standard setting of $P2_1/a$. This showed that the sizes and shapes of unit cells of α'-(2) and α-(2), neglecting the *C*-centring of the former, and allowing for the effects of thermal expansion, are virtually identical. Furthermore, the addition of a crystallographic mirror plane at $y = 0$ to the structure of α-(2) at 150 K converts the $P2_1/a$ space group into its supergroup $C2/m$, as required for the room-temperature form of α'-(2). As is evident in Figs. 1 and 2, the anisotropic displacement ellipsoids are significantly larger for α'-(2) than for α-(2), but more importantly those of the S atoms are much larger in α'-(2) than in α-(2). A refinement for α'-(2) in which S3 was modelled as being disordered about the mirror plane led to essentially identical residuals.

In Table 2 it is noticeable that despite the larger cell of α'-(2) the bond distances, especially those involving S atoms, are generally shorter than those observed in α-(2), although no corrections for libration have been made. The bond-length differences are not large and not individually statistically significant, but they are consistently in the same sense. It is concluded that the symmetry of the molecule as depicted in α'-(2) arises from the superimposition of the images of the two enantiomeric conformations of α-(2). There is a concern that the comparatively high (*C*-centred) symmetry of α'-(2) has come about due to a failure to measure weak reflections with $h + k$ odd. However, a further data collection using a Bruker SMART 1000 area detector system and paying particular attention to this very point, completely confirmed the choice of the *C*-centred cell for α'-(2). There was little evidence for diffuse scattering in the CCD data frames.

The interchange between U_R and U_L conformations involves rotations about the S3—C3 and S5—C8 bonds. The S1 atom in U_R (or U_L) is displaced from its position in the C_v symmetric U conformation by ∼1/4 of the diameter of a benzene ring, *i.e.* approximately 0.7 Å. The C2—S3—C3—C4 and C7—S5—C8—C9 dihedral angles provide another estimate of the effect. In the U conformation of α'-(2) these angles are equal in magnitude [8.7 (5)°], but opposite in sign. In the conformations in α-(2) these angles are −17.10 (13) and −33.34 (13)°, respectively; the average change in these angles over the six occurrences in the three molecules with U_R/U_L conformations being 25.47°. In the transformation from U_R to U_L, the S1 displacement and the change in dihedral angle each has twice the values given above. Both these measures tend to exaggerate the degree of distortion because the change in value has a component arising from the displacement of the central atoms, especially S3 and S5.

The fact that the change in conformation from U_R to U_L requires no more than bond twisting is further evidence in support of molecular libration as the probable cause of the order–disorder relationship between α'-(2) and α-(2), which in turn is the justification for their α' and α phase designations.

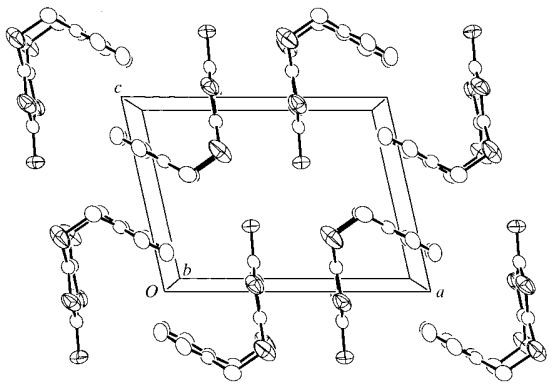


Figure 5
Packing diagram for α'-(2).

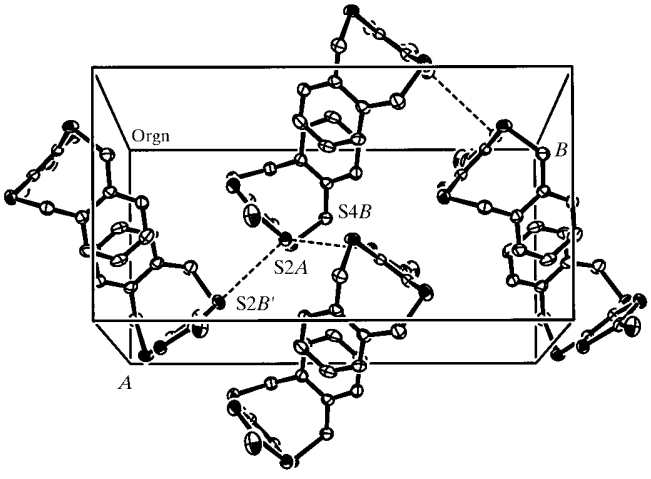


Figure 6
Layer of molecules in β-(2). Dashed lines indicate short S···S contacts (see text).

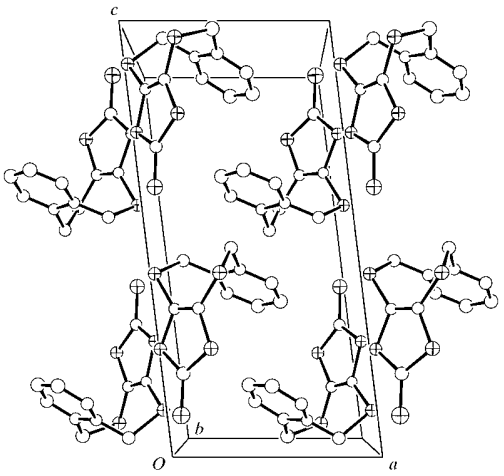


Figure 7
Packing diagram for β-(2).

The packing of the molecules of α'-(2) and α-(2) in their unit cells is essentially the same. As shown in Fig. 4 the molecules of α-(2), when viewed along **b**, lie edge-on, forming layers centred on $y = 0$ and $y = 1/2$. Within the layers, both the benzene rings and the dmit moieties occur face-to-face in centrosymmetrically related pairs. In α'-(2) (Fig. 5), the perpendicular distances between benzene ring pairs, centred on 0,0,0 and 0,0,1/2, are 3.276 (14) and 3.544 (8) Å, respectively, while the perpendicular distance between the dmit planes defined by the C2═C7 double bond and its attached S atoms is 3.748 (4) Å. In α-(2), the corresponding distances are 2.959 (5), 3.554 (4) and 3.630 (6) Å.

Despite both molecules in the asymmetric unit having conformations essentially the same as in α-(2), the packing of the molecules in the β-(2) phase is completely different. As shown in Fig. 6, the molecules occur in *AB* pairs, arranged in a herringbone-type array, to form a layer parallel to (001) and with z in the range 0–1/2. Short S···S contacts of 3.3259 (9) Å [S2*A*···S4*B*] and 3.5568 (9) Å (S2*A*···S2*B*′) occur between the pairs of molecules. A second layer, related to the first by the operation of a crystallographic centre of symmetry, completes the unit-cell contents. Centrosymmetric face-to-face benzene ring contacts occur from one layer to the next at $z = 0$ and 1/2 (see Fig. 7) at perpendicular distances of 3.537 (3) [*AA* pairs] and 3.574 (3) Å [*BB* pairs]. In contrast, there are no face-to-face contacts of the dmit moieties within the sum of the van der Waals' radii for two S atoms (3.70 Å; Huheey *et al.*, 1993). It is the gross difference in molecular packing which prompts the designation of β-(2) as a true polymorph.

The molecules of β-(2) can be regarded as forming a column comprising non-centrosymmetric *AB* molecule pairs, centred on z = 1/4 and 3/4, each pair being related to the next by a crystallographic centre of symmetry at $z = 0$ or 1/2. Similar columns are also present in α-(2) and α'-(2), but they are less obvious: here the equivalent of the *AB* molecule pairs of β-(2) are centrosymmetric, the columns are propagated in the *a* direction and arranged side by side in the *c* direction in such a way as to create centrosymmetrically related pairs of dmit moieties and so complete the layers.

We thank the EPSRC National Crystallography Service (University of Southampton) for data collections.

References

Allen, F. H. & Kennard, O. (1993). *Chem. Des. Autom. News*, **8**, 31–37.
Blessing, R. H. (1997). *J. Appl. Cryst.* **30**, 421–426.
Culot, P., Dive, G., Nguyen, V. H. & Ghuysen, J. M. (1990). *Theoret. Chim. Acta*, **82**, 189–205.
Desiraju, G. R. (1995). *Angew. Chem. Int. Ed. Engl.* **34**, 2311–2327.
Fletcher, D. A., McMeeking, R. F. & Parkin, D. (1995). *J. Chem. Inf. Comput. Sci.* **36**, 746–749.
Goldenberg, L. M. & Lyubovskaya, R. N. (1986). *Khim. Geterotsikl. Soed.* **6**, 855–856.
Hooft, R. W. W. (1998). *Collect.* Enraf–Nonius, Delft, The Netherlands.
Howie, R. A. (1980) *RDNIC.* University of Aberdeen, Scotland.
Huheey, J. E., Keiter, E. A. & Keiter, R. L. (1993). *Inorganic Chemistry, Principles of Structure and Reactivity*, 4th ed. New York: HarperCollins.
Hunter, C. A. (1994). *Chem. Soc. Rev.* **23**, 101–109.
Kumar, S. K., Singh, H. B., Jasinski, J. P., Paight, E. S. & Butcher, R. J. (1991). *J. Chem. Soc. Perkin Trans 1*, pp. 3341–3347.
Mohamadi, F., Richards, N. G. J., Guida, W. C., Liskamp, R., Lipton, M., Caulfield, C., Chang, G., Hendrickson, T. & Still, W. C. (1990). *J. Comput. Chem.* **11**, 440–467.
Nicolet Instrument Corporation (1980). *P*3 *Software.* Nicolet Instrument Corporation, Madison, Wisconsin, USA.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sheldrick, G. M. (1990). *Acta Cryst.* A**46**, 467–473.
Sheldrick, G. M. (1997). *SHELXL*97. University of Göttingen, Germany.
Still, W. C., Tempczyk, A., Hawley, R. C. & Hendrickson, T. (1990). *J. Am. Chem. Soc.* **112**, 6127–6129.
Wang, C., Batsanov, A. S., Bryce, M. R. & Howard, J. A. K. (1998). *Synthesis*, pp. 1615–1618.

J. Membrane Biol. 176, 119–131 (2000)
DOI: 10.1007/s00232001078

The Journal of
Membrane
Biology

Analysis of the Structure and Electrophysiological Actions of Halitoxins: 1,3 Alkyl-pyridinium Salts from *Callyspongia ridleyi*

R.H. Scott[1], A.D. Whyment[1], A. Foster[1], K.H. Gordon[2]*, B.F. Milne[2], M. Jaspars[2]
[1]Department of Biomedical Sciences, Institute of Medical Sciences, Aberdeen University, Foresterhill, Aberdeen AB25 2ZD, UK
[2]Marine Natural Products Laboratory, Department of Chemistry, University of Aberdeen, Old Aberdeen, AB24 3UE, UK

Received: 23 December 1999/Revised: 3 April 2000

Abstract. We have chemically characterized a preparation of halitoxins, (1,3 alkyl-pyridinium salts) isolated from the marine sponge *Callyspongia ridleyi*. At concentrations of 50 and 5 μg/ml the halitoxin preparation caused irreversible membrane potential depolarization, decreased input resistance and inhibited evoked action potentials when applied to cultured dorsal root ganglion neurones. Under whole cell voltage clamp the halitoxins produced an increase in cation conductance that was attenuated by replacing sodium with N-methyl-D-glucamine. Fura-2 fluorescence ratiometric calcium imaging was used to directly measure calcium flux into neurones after exposure to halitoxins. Calcium influx, evoked by the halitoxins, persisted when the neurones were bathed in medium containing the voltage-activated calcium channel antagonists cadmium and nickel. Experiments on undifferentiated F-11 cells showed little or no calcium influx in response to depolarizing concentrations of potassium and indicated that halitoxins evoked massive calcium influx in the absence of voltage-activated calcium channels. The halitoxins also produced transient increases in intracellular calcium when F-11 cells were bathed in calcium-free medium suggesting that the toxins could release calcium from intracellular stores. The pore-forming action of the halitoxins was identified when the toxins were applied to artificial lipid bilayers composed of phosphatidylcholine and cholesterol. Halitoxins evoked channel-like activity in the lipid bilayers, with estimated unitary conductances of between 145pS and 2280pS, possibly indicating that distinct channels could be produced by the different components in the preparation of halitoxins.

Key words: Calcium permeant ion channel — Pore former — Halitoxin — Sensory neurone — Lipid bilayer

Introduction

Recent years have seen the growth of a body of literature relating to a number of secondary metabolites with a common 3-alkyl pyridinium or 3-alkyl piperidine moiety isolated from a variety of genera in the phylum *Porifera* (the sponges) (Andersen, Van Soest & Kong, 1996). Examples of these are the halicyclamines (Jaspars et al., 1994) and the manzamines (Crews et al., 1994). These alkaloids all display marked biological activity and may function as sponge chemical defenses against predation and invasion by microorganisms. Therefore these compounds are of interest for their potential therapeutic properties and as tools for research. Studies aimed at the isolation and characterization of these compounds have elucidated much about their structure and composition. During our work we focused on the structure and electrophysiological properties of one set of these compounds, the 1,3-alkylpyridinium salts (1,3-APS) examples of which are the halitoxins (Fig. 1.1; Schmitz, Hollenbeak & Campbell, 1978) and the amphitoxins (Fig. 1.2; Albrizio et al., 1995). Varied biological activity has been reported for these compounds from cytotoxicity (Schmitz et al., 1978) through to epidermal growth factor inhibition (Davies-Coleman et al., 1993), neurotoxicity (Berlinck et al., 1996) and anticholinesterase activity (Sepcic et al., 1997). The halitoxin compounds have been isolated from a wide number of Haplosclerid genera: *Haliclona erina; Haliclona rubens; Haliclona viridis* (Schmitz et al., 1978); *Amphimedon viridis* (for-

**Present address:* University Chemical Laboratory, Lensfield Road, Cambridge, CB2 1EW

Correspondence to: R.H. Scott

merly *Haliclona viridis,* Berlinck et al., 1996); *Amphimedon compressa* (Albrizio et al., 1995); *Callyspongia fibrosa* (Schmitz et al., 1987) and *Reniera sarai* (Sepcic et al., 1997) suggesting they may have common functional properties as poriferan defense system. 1,3-APS occur as high molecular weight oligomers ranging from 1 KDa to greater than 25 KDa, and it is most likely that they exist as linear oligomers, with varying lengths of aliphatic chains linking the pyridine units. Little is known about the supramolecular structuring of these compounds except for the determination of average hydrodynamic radii of high weight oligomers (Sepcic et al., 1997).

Relatively little work has been carried out on the biological actions of halitoxins, but several previous studies suggest that an electrophysiological investigation would be appropriate. Firstly, Baslow and Turlapaty (1969) identified antitumor activity of halitoxins and at higher doses neurotoxic actions consisting of tremors, convulsions and subsequent paralysis. Secondly, halitoxin has been found to specifically inhibit potassium conductances in frog muscle (Sevcik et al., 1986) and squid axon (Sevcik et al., 1994) preparations. Thirdly, further studies have shown that successive application of halitoxin to a crab nerve preparation depolarizes membrane potential and suppresses action potentials in a dose-dependent manner. Additionally, halitoxin produced dose-dependent lysis of sea urchin eggs. These actions of halitoxins may be indicative of these molecules having nonspecific actions on biological membranes (Berlinck et al., 1996). We have therefore determined the chemical composition of our halitoxin fraction and investigated its actions on cultured DRG neurones, undifferentiated F-11 cells (a mouse dorsal root ganglion neurone × N18TG2 neuroblastoma hybridoma, Platika et al., 1985) and a lipid bilayer preparation.

Materials and Methods

Isolation and Structure Determination

During this study a collection of *Callyspongia ridleyi* from Papua New Guinea was investigated. Collection, solvent extraction and partition were carried out using the standard protocol described earlier (Jaspars et al., 1994). The butanol extract contained the bulk of the 1,3-APS compounds as indicated by nuclear magnetic resonance spectroscopy (NMR). This was applied to a lipophilic Sephadex LH-20 size exclusion column and several fractions were collected. For further investigations we chose the fraction whose molecular weight was shown to be in the range of 5–6 KDa. The molecular weights were determined by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS).

Calculation of charge distribution in 1-methyl-3-picolinium and 1-methyl pyridinium cations was carried out. This was done in order to gain some insight into the distribution of the +1 charge formally associated with the quaternized nitrogen atom in the 1,3-alkyl pyridinium subunit characteristic of the halitoxins. The model systems 1-methyl-3-picolinium (Fig. 1.4) and 1-methyl pyridinium (Fig. 1.5) were chosen. *Ab-initio* calculations were performed at the MP2/6-31G(d)//RHF/6-31G(d) level using the GAMESS-US quantum-chemical software package (Schmidt et al., 1993) and the partial atomic charges derived from the resulting nuclear geometry and electronic distribution.

Cell Cultures

Primary cultures of DRG neurones were prepared following enzymatic and mechanical dissociation of dorsal root ganglia from decapitated 2-day old Wistar rats. The sensory neurones were plated on laminin-polyornithine coated coverslips and bathed in F14 culture medium (Imperial Laboratories) supplemented with 10% horse serum (Gibco), penicillin (5000 IU/ml), streptomycin (5000 mg/ml), $NaHCO_3$ (14 mM) and nerve growth factor (20 ng/ml). The cultures were maintained for up to 3 weeks at 37°C in humidified air with 5% CO_2, and re-fed with fresh culture medium every 5–7 days.

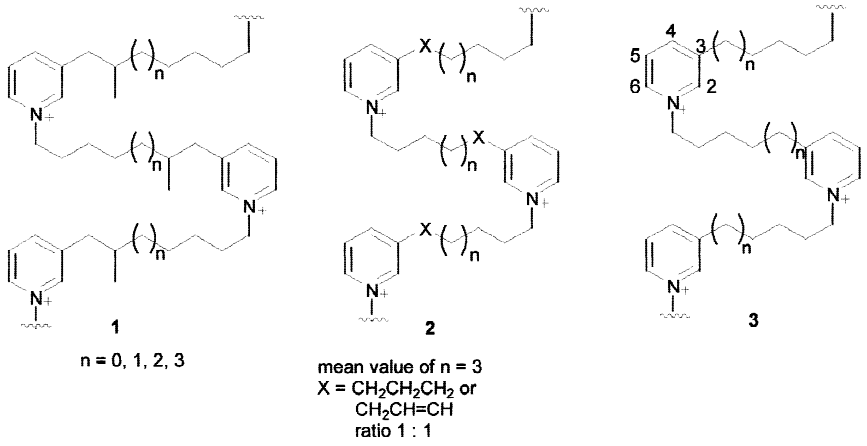
Undifferentiated F-11 cells from Dr. M.C. Fishman (Massachusetts General Hospital, Boston, MA) were grown as a monolayer on glass coverslips in 30 mm dishes. The cells were bathed in Ham's F-12 culture medium containing 2 mM L-glutamine supplemented with 15% Hyclone fetal bovine serum, HAT supplement (100 μM hypoxanthine/400 nM aminopterin/16 μM thymidine), 100 units/ml penicillin and 100 μg/ml streptomycin. The cultures were maintained at 37°C in humidified air with 5% CO_2, and passed twice a week using mechanical cell dissociation.

Electrophysiology

Experiments were conducted at room temperature (approximately 23°C) on DRG neurones that had been in culture for at least 2 days. The whole-cell recording technique (Hamill et al., 1981) was used to study the actions of halitoxins on membrane potential, evoked action potentials and input resistance. Additionally, under voltage-clamp halitoxin-evoked whole cell currents were investigated. Whole cell recordings were made using either an Axopatch-1D amplifier or an Axoclamp-2A switching voltage-clamp amplifier. Under voltage clamp, seventy to eighty percent series resistance compensation was applied and the Axoclamp-2A was operated at a sampling rate of 15–20 kHz. Low resistance (5–12 MΩ) borosilicate glass patch pipettes were fabricated using a Kopf model 730, needle/pipette puller. The neurones were bathed in a variety of different extracellular solutions. The standard NaCl-based extracellular solution contained in mM: NaCl, 130; KCl, 3.0; $CaCl_2$, 2.0; $MgCl_2$, 0.6; $NaHCO_3$ 1.0, HEPES 10.0, glucose 5.0. The N-methyl-D-glucamine (NMDG)-based extracellular solution contained in mM: NMDG, 166; $CaCl_2$, 2.0; $NaHCO_3$ 1.0, HEPES 10.0, glucose 4.0, tetrodotoxin (TTx, Alomone labs) 0.0025. The NMDG (Ca^{2+}-free)-based extracellular solution contained in mM: NMDG, 166; $NaHCO_3$ 1.0, HEPES 10.0, glucose 4.0, TTx 0.0025. For all the extracellular solutions the pH and osmolarity were adjusted to 7.4 and 310–320 mOsmol/l with NaOH and sucrose, respectively. The patch pipette solution contained in mM: KCl, 140; EGTA, 5; $CaCl_2$, 0.1; $MgCl_2$, 2.0; HEPES, 10.0; ATP, 2.0; and the pH and osmolarity were adjusted to 7.2 with Tris and 310–315 mOsmol/l with sucrose. After entering the whole cell recording configuration, neurones were allowed to equilibrate for 5 min before electrophysiological measurements were made.

For voltage clamp recordings, neurones were held at –90 mV and linear current-voltage relationships were generated with 100 msec voltage step commands to potentials between –170 and –60 mV under control conditions and –170 and +20 mV after halitoxin application.

Halitoxins were applied to the extracellular environment by low pressure ejection via a blunt micropipette (tip diameter about 10 μm) positioned approximately 100 μm from the neurone being recorded. The cells were maintained in a bath and were not continually perfused; drug concentrations declined after pressure ejection as a result of diffusion.



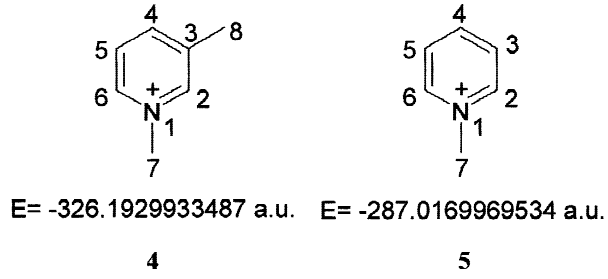


Fig. 1. Structures of halitoxins and related compounds. (1) halitoxins, (2) amphitoxins and (3) 1,3 alkyl-pyridinium salts with saturated alkyl chains, investigated in this study. Structures of (4) 1-methyl-3-picolinium and (5) 1-methyl pyridinium cations used in the charge distribution calculations.

The electrophysiological data were stored on digital audiotape (DAT) using a DTR-1200 DAT recorder (Biologic) and subsequently analyzed using Cambridge Electronic Design voltage clamp software (version 6). For monitoring changes in membrane potential or holding current continuous records were obtained on a chart recorder (Gould 2200s pen recorder). All data are given as Mean ± SEM and statistical significance was determined, using the Student's two-tailed *t* test, paired or independent where appropriate and *P* values are reported in the text.

Intracellular Calcium Imaging

Cultured DRG neurones and F-11 cells were incubated for 1 hr in NaCl-based extracellular solution containing 10 μM fura-2AM (Sigma, 1 mM stock in dimethylformamide). The cells were then washed for 20 min to remove the extracellular fura-2AM and to allow cytoplasmic de-esterification of the Ca^{2+}-sensitive fluorescent dye. The cells were constantly perfused (1–2 ml/min) and viewed under an inverted Olympus BX50WI microscope with a KAI-1001 S/N 5B7890-4201 Olympus camera attached. The fluorescence ratiometric images from data obtained at excitation wavelengths of 340 nm and 380 nm were viewed and analyzed using OraCal pro, Merlin morphometry temporal mode (Life Sciences resources, version 1.20). All experiments were conducted at room temperature and data are expressed as Means ± SEM.

Lipid Bilayer and Pipette Dipping

The pipette dipping method (Coronado & Latorre, 1983) was used to investigate the actions of halitoxins on artificial planar lipid bilayers. All our recordings were made under voltage-clamp conditions using an Axopatch-1D amplifier and 10 MΩ borosilicate glass patch pipettes dipped in "Sigma-coat" and filled with the KCl-based patch pipette solution. The standard NaCl-based solution was in the bath. Lipid bilayers were formed using a monolayer of phosphatidylcholine (Sigma, type 2s from Soy Bean) and cholesterol (Sigma, grade 1 from Porcine liver), at a molar ratio of 9:1 in HPLC grade *n*-pentane. The tip of the patch pipette was placed in NaCl-based extracellular solution contained in a 35 mm dish and 10 to 30 μl of a 1 mg/ml lipid/*n*-pentane solution was applied to the solution surface. After 2 to 5 min the pentane evaporated from the surface of the monolayer. The patch pipette was gently removed from the solution and then replaced into the solution to form a planar lipid bilayer across the tip of the pipette. The seal resistance of the lipid bilayers ranged from 6 to 32 GΩ. The bilayer seal resistances were stable for 10 to 20 min before halitoxin 50 μg/ml was applied by pressure ejection from a blunt patch pipette.

Results

Chemical Data

By comparison to literature NMR data the structure was identified, *see* Fig. 1.3, with saturated alkyl chains. We were interested in defining the size and size distribution of the alkyl chains in more detail, as it was not clear whether the number *n* in 3 varied within each chain or only between different chains. We used electrospray mass spectrometry (ESI-MS) for this, and the distribution of 1,3-APS monomers is shown in Table 1. It can be seen that there is a preference for 8 and 9 carbon alkyl connecting chains ($n = 4, 5$ in Fig. 1.3), and that a 5 carbon chain is the shortest observed ($n = 1$ in Fig. 1.3).

Analysis of the dimer units from the same ESI mass spectrum indicated that dimers could be formed from various combinations of monomer units (Table 2). The fragmentation process was assumed to be that described

Table 1. Distribution of monomers by ESI-MS in 1,3-APS isolated from *C. ridleyi*

Monomer number	Mass fragment (*m/z*)	Monomer unit (inc pyr)	Number of carbons in alkyl chain (value of *n* in Fig. 1.2)	Relative abundance in MS (%)
1	148	$C_{10}H_{14}N^+$	5 (1)	10
2	162	$C_{11}H_{16}N^+$	6 (2)	15
3	176	$C_{12}H_{18}N^+$	7 (3)	13
4	190	$C_{13}H_{20}N^+$	8 (4)	85
5	204	$C_{14}H_{22}N^+$	9 (5)	100
6	218	$C_{15}H_{24}N^+$	10 (6)	36
7	232	$C_{16}H_{26}N^+$	11 (7)	5

Table 2. Dimers observed by ESI-MS and their molecular weights

Monomer	1	2	3	4	5	6	7
1	yes	yes	yes	yes	yes	yes	yes
	295	309	323	337	351	365	379
2		yes	yes	yes	yes	yes	yes
		323	337	351	365	379	393
3			yes	yes	yes	yes	no
			351	365	379	393	
4				yes	yes	no	no
				379	393		
5					no	no	no
6						no	no
7							no

by Davies-Coleman et al., 1993. If each 1,3-APS polymer chain was composed of only one monomer type then we would expect only homodimers to be observed (e.g., 1–1, 2–2, 3–3, etc). However, we see many peaks in the spectrum attributable only to heterodimers (e.g., 1–2, 3–5 etc), suggesting that different monomers are present within each 1,3-APS polymer chain. It is impossible to differentiate between a dimer consisting of say, monomers 2 and 3, and one containing monomers 1 and 4, as these have the same molecular mass, though it is likely that both are present in the mixture. Even segregation into polymers containing only chains with an even or odd number of carbon chains can be ruled out by the existence of such dimers as 1–4 and 2–3 (both *m/z* 337). It is interesting to note that no higher molecular weight dimers are present in the mixture, although these are theoretically possible (e.g. 5–5, 6–7, 7–7, etc). The most abundant peaks are at *m/z* 379 (48%, 4–4, 3–5, 2–6, 1–7 dimers), 295 (30%, 1–1 dimer) and 393 (28%, 5–4, 3–6, 2–7 dimers), which is in keeping with the observed abundance of the monomers, except for the peak at *m/z* 295. We do not know why there is this exceptional peak at *m/z* 295, it could be due to the mass spectrometry conditions and 295 being stable to cleavage, but perhaps it relates to the sponge's biochemistry and that the organisms do not produce the monomer at 148.

Besides monomers and dimers only a series of hexamers was observed in the ESI mass spectrum at *m/z* 1100–1300. Analysis of isotope patterns showed each of these to contain 5 Cl^- to counterbalance the charge on 6 pyridine N^+. One example is the cluster at *m/z* 1185.8 which might be composed of $[(C_{10}H_{14}N)_3 + (C_{11}H_{16}N) + (C_{12}H_{18}N) + (C_{16}H_{26}N)]Cl_5 = C_{69}H_{102}N_6Cl_5$ (MW 1189.7). The system would need to lose four protons which can occur readily in the mass spectrometer during the fragmentation process (Davies-Coleman et al., 1993). A similar analysis can be carried out for the other clusters observed. This again indicates that the 1,3-APS polymer chain is made up of a random sequence of monomers.

The MALDI-TOF MS spectrum showed peaks at *m/z* 4713 (38), 4827 (8), 4998 (100), 5510 (8). The base peak at *m/z* 4998 translates to roughly 19–27 monomer units (*see* Table 1) in the chain if the chloride ions are included.

NMR Data for the Fraction Which is ~5KDa from Chromatography

δ_H (250 MHz, D_2O, HDO ref. at 4.80 ppm) 8.7 ppm, 2H, s (pyr H2 and pyr H6); 8.4 ppm, 1H, d, $J = 7.3$ Hz (pyr H4); 8.0 ppm, 1H, t, $J = 7.0$ Hz (pyr H5); 4.5 ppm, 2H, t, $J = 7.3$ Hz (N^+-CH_2); 3.2 ppm, 2H, t, $J = 7.6$ Hz (pyr-CH_2); 2.8 ppm, 2H, m (N^+-CH_2-CH_2); 2.0-1.0 ppm, m, (alkyl CH_2). δ_C (62.9 MHz, D_2O) 145 ppm (pyr C4); 144 ppm (pyr C3); 143 ppm (pyr C2); 142 ppm (pyr C6); 128 ppm (pyr C5); 62 ppm (N^+-CH_2); 32 ppm, 31 ppm, 30 ppm, 28 ppm (several peaks), 26 ppm (all alkyl CH_2).

Relative abundance of dimer peaks in ESI-MS (based on *m/z* 204 = 100%) 295 (30); 309 (22); 323 (13); 337 (7); 351 (8); 365 (21); 379 (48); 393 (28).

Isotope pattern for cluster beginning at *m/z* 1185.8: 1185.8 (50); 1186.9 (33), 1187.8 (100); 1188.7 (63); 1189.8 (70); 1190.8 (50); 1191.9 (33); 1192.8 (17). Calculated for $C_{69}H_{98}N_6Cl_5$: 1185.6 (52); 1186.6 (42); 1187.6 (100); 1188.6 (72); 1189.6 (80); 1190.6 (50); 1191.6 (36); 1192.6 (19). Other hexamer clusters appeared beginning at *m/z* 1157.9 (20); 1171.9 (100); 1185.8 (70); 1299.9 (40); 1213.9 (18); 1228.0 (8); 1242.0 (6); 1256.1 (3).

In summary the structures of the compounds studied in this project are given in Fig. 1.3 with *n* ranging from one to seven and the number of monomers ranging from 19 to 27.

Calculation of Charge Distribution in 1-Methyl-3-Picolinium and 1-Methyl Pyridinium Cations

From Table 3 it can be seen that the formal unit positive charge associated with the nitrogen is in fact distributed

over carbons 2 and 6 and all of the protons in the molecule. The nitrogen itself carries a negative charge and would not therefore be expected to interact favorably with anionic species as might be inferred from its formal charge state. As the highest concentration of positive charge occurs on the hydrogens it appears that a less localized electrostatic interaction with the "edge" of the pyridine ring would be more likely.

Comparison with the molecule illustrated in Figs. 1.4 and 5 shows that there is a significant rearrangement of charge on the heavy atoms introduced by the desymmetrization of the aromatic centers. However, the resulting charges on the ring protons are only very slightly effected, as are the charges associated with the nitrogen and its adjacent carbons.

Table 3. Mulliken partial atomic charges obtained from MP2/6-31G(d)//RHF/6-31G(d) calculation on Fig. 1.4 and Fig. 1.5.

	1-Methyl-3-picolinium (4)	1-Methyl pyridinium (5)
ATOM	Charge (units = *e*)	
N1	–0.5849	–0.5863
C2	+0.1350	+0.1607
C3	–0.0493	–0.2698
C4	–0.0943	–0.0869
C5	–0.2738	–0.2721
C6	+0.1599	+0.1652
C7	–0.5182	–0.3340
C8	–0.3331	######
H2	+0.2878	+0.2980
H3	######	+0.2942
H4	+0.2864	+0.2960
H5	+0.2910	+0.2940
H6	+0.2961	+0.2988
H7a*	+0.2389	+0.2404
H7b*	+0.2486	+0.2508
H7c*	+0.2486	+0.2508
H8a*	+0.2099	######
H8b*	+0.2257	######
H8c*	+0.2257	######
TOTAL =	+1.0000	+1.0000

Atomic units used ($1e = 1.602 \times 10^{-19}$ C).
* Note: differences in the charges on methyl protons are due to the fact that one of the three hydrogens in each group lies in the plane of the heavy atoms and hence is subject to a different electrostatic environment than those lying above and below this plane.

Actions of Halitoxins on the Electrophysiological Properties of Cultured DRG Neurones

In this study, three different doses of 0.5, 5.0 and 50 μg/ml halitoxin fraction were used. Given a mean estimated molecular weight of 5 KDa for the toxin preparation the approximate concentrations of halitoxins were 0.1, 1.0 and 10 μM. Under control conditions the resting membrane potential of the cultured DRG neurones was –60 ± 2 mV ($n = 27$). Application of halitoxins for 20 sec caused an apparent dose-dependent depolarization to –26 ± 5 mV ($n = 15$, $P < 0.001$) and –7 ± 1 mV ($n = 9$, $P < 0.001$), for 5 μg/ml and 50 μg/ml halitoxins, respectively. Application of 0.5 μg/ml halitoxins for 20 sec gave no significant change in resting membrane potential. The depolarizations evoked by the halitoxins were associated with significant reductions in the mean input resistances of the DRG neurones. Under control conditions the mean input resistance was 356 ± 42 MΩ ($n = 12$). In the presence of 0.5, 5 and 50 μg/ml halitoxins the mean input resistance values were 283 ± 34 MΩ ($n = 3$, NS), 71 ± 23 MΩ ($n = 6$, $P < 0.05$) and 27 ± 12 MΩ ($n = 3$, $P < 0.02$), respectively. The membrane potential and input resistance did not significantly recover after 20 sec application of the halitoxins, even if the cells were left for 20 min or more (Fig. 2*A*). In a further set of experiments current was injected into the cells to compensate for the depolarizations produced by the halitoxins. This was done so that input resistance and evoked action potentials could be consistently measured from a potential of –70 mV under control conditions and after the halitoxins had been applied. Even when corrections were made for changes in membrane potential significant reductions in input resistance were measured following 20 sec application of 5 and 50 μg/ml halitoxins (Fig. 2*B*). The effects of the halitoxins on input resistance were seen over a wide voltage range (Fig. 2*C*) and were coupled with an inability to activate an action potential even when supramaximal stimuli were applied (Fig. 2*D*). If 5 μg/ml of halitoxins were applied for two periods of about 20 sec then the toxins produced additional changes in membrane potential and input resistance, suggesting that steady-state responses were not achieved using this protocol (Fig. 3*A*). This could be due to the effects of the halitoxins involving a mechanism that could not be saturated, rather than a conventional receptor-ligand interaction. The changes in membrane potential were partially but significantly reversed by perfusion of NMDG-based extracellular solution (Fig. 3*A* and *B*). However, NMDG-based extracellular solution failed to significantly reverse the effects of the halitoxins on input resistance, with mean values of 20 ± 4 MΩ ($n = 3$) and 56 ± 15 MΩ ($n = 4$) in NaCl-based and NMDG-based extracellular solutions, respectively. These results could have been due to the presence of 2.5 μM TTx in the NMDG-based extracellular solution blocking voltage-activated Na^+ channels. To investigate this possibility, responses to halitoxins (50 and 5 μg/ml; n = 3) were evoked in NaCl-based extracellular solution before TTx was applied in an attempt to reverse part of the toxins actions. Application of TTx (2.5 μM) did not cause any change in either the membrane potential or input resistance after treatment with halitoxins (Fig. 3*C*), which

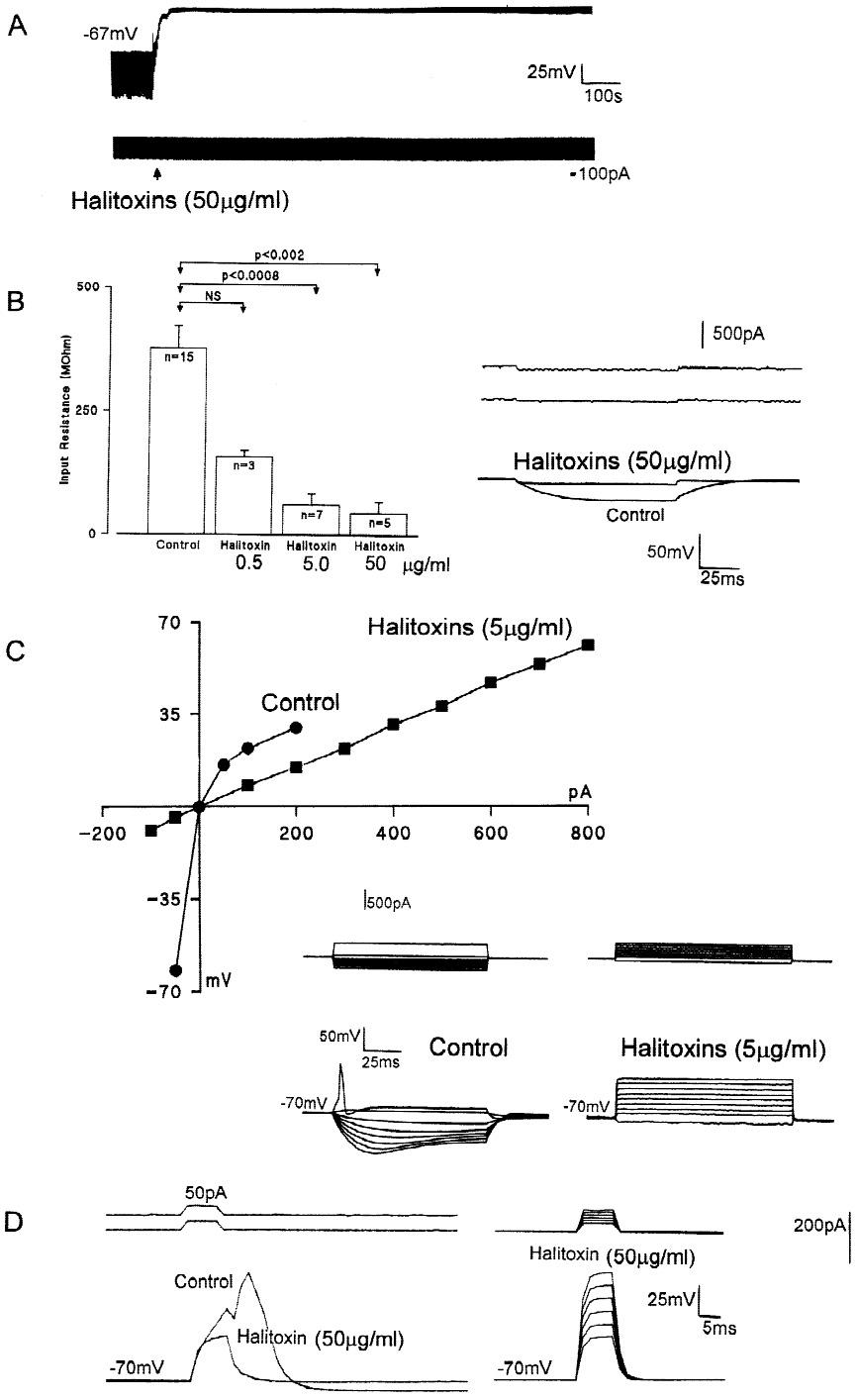


Fig. 2. Electrophysiological actions of halitoxins on cultured DRG neurones. (*A*) Record showing depolarization and change in input resistance produced by a 20 sec application of halitoxins (50 μg/ml). Electrotonic potentials were evoked by −100 pA hyperpolarizing current commands of 100 msec durations applied every 3 sec. (*B*) Bar chart showing the mean input resistances of neurones under control conditions and after exposure to varying concentrations of halitoxins. All measurements were made from a clamp potential of −70 mV. The inset traces show current commands and electrotonic potentials recorded in the absence (Control) and presence of halitoxins. (*C*) Current/voltage relationships obtained under control conditions (filled circles) and after 20 sec application of halitoxins (filled squares). Inset records show example current commands and potentials recorded from a clamp potential of −70 mV, under control conditions and after application of halitoxins. (*D*) Records show the abolition of an evoked action potential (control) by halitoxins. The right hand records show that this effect could not be overcome by supramaximal depolarizing current commands.

suggests that the effect of NMDG-based extracellular solution was due to lower permeability of the large cation compared with Na^+.

Under voltage-clamp conditions 50 μg/ml of halitoxins produced such large inward currents it was not possible to clamp the neurones and so further studies were carried out with 5 μg/ml. The neurones were voltage clamped at −90 mV and current/voltage relationships generated under control conditions and after application of the toxins. Under control conditions neurones were not depolarized to potentials positive to −60 mV, because the linear, (ohmic) current-voltage relationship would be contaminated with voltage-activated conductances. This was not a problem after application of the halitoxins that evoked large inward currents with a mean value of −2.65 ± 0.51 nA ($n = 5$) from a holding potential of −90 mV, and appeared to abolish voltage-activated currents. In NaCl-based extracellular medium the mean conductance and reversal potential for the current evoked by halitoxins were 30 ± 6 μS and −1 ± 4 mV, respectively ($n = 5$, Fig. 4*A*). In NMDG-based extracellular solution (with Ca^{2+} present), the current evoked by halitoxins had a reduced conductance of 12 ± 4 μS ($n = 5, P < 0.02$) and a more hyperpolarized reversal potential of −40 ± 9 mV

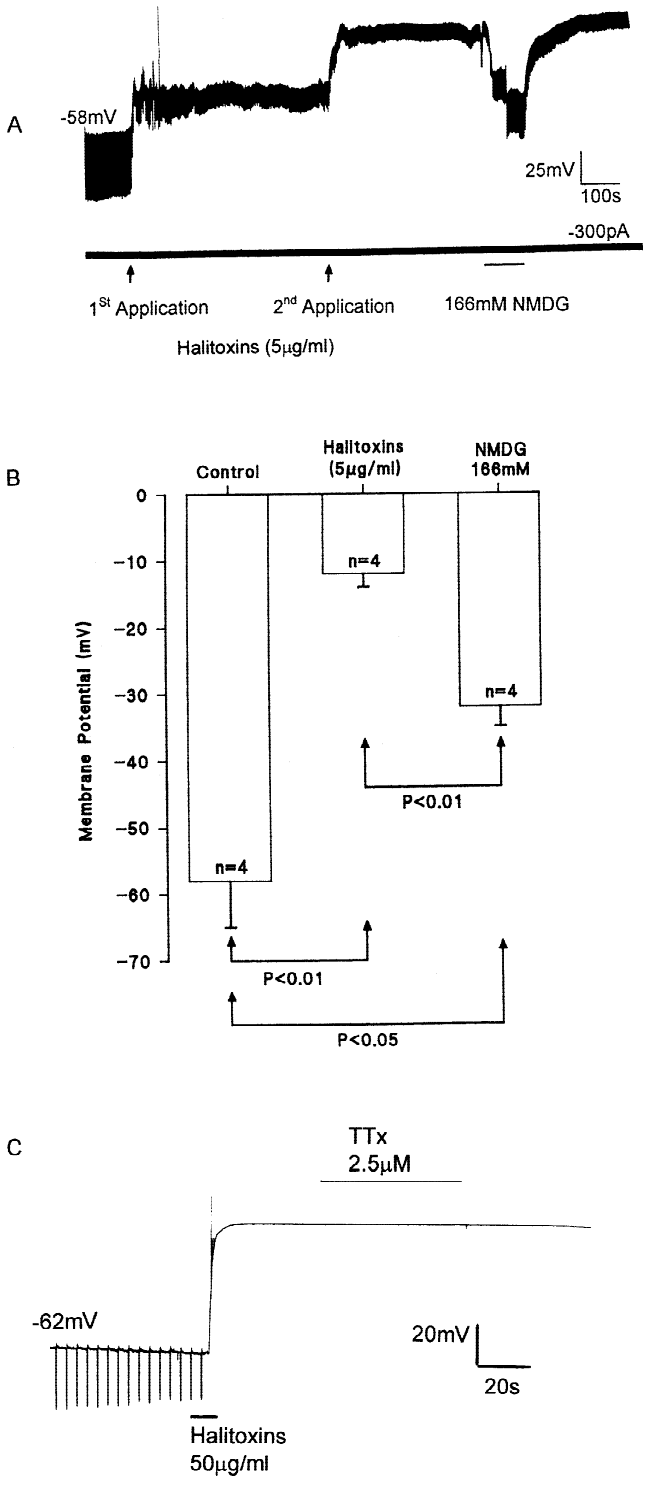


Fig. 3. NMDG-based recording medium attenuated the effects of halitoxins on membrane potential. (*A*) A recording, showing the effects of two applications of halitoxins on membrane potential and input resistance. Electrotonic potentials were evoked every 3 sec by 100 sec hyperpolarizing current commands (–300 pA). Application of NMDG-based extracellular solution caused a partial repolarization of membrane potential. (*B*) Bar chart showing the mean membrane potentials for four neurones under control conditions, after application of halitoxins in NaCl-based extracellular solution and after the response to the halitoxins but during application of NMDG-based extracellular solution. (*C*) A recording showing that 2.5 μM TTx in NaCl-based extracellular solution failed to reverse the actions of halitoxins on membrane potential and input resistance.

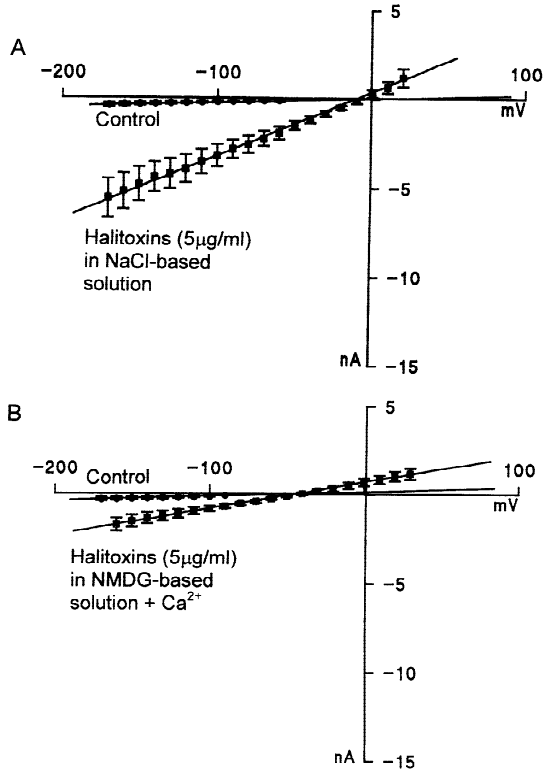


Fig. 4. Current/voltage relationships under voltage clamp in the presence of NaCl-based extracellular solution (*A*) and in the presence of NMDG-based extracellular solution (*B*). In both cases the neurones were voltage-clamped at a holding potential of –90 mV. Depolarizing and hyperpolarizing voltage step commands were applied for 100 msec to generate current-voltage relationships before (Control) and after application of halitoxins. Mean data ± SEM are presented, and $n = 5$ for each experiment.

($n = 5$, $P < 0.05$, Fig. 4*B*). These results suggested significant components of the conductances were due to Ca^{2+} and/or NMDG. Experiments carried out in NMDG-Ca^{2+} free extracellular solution gave anomalous results. No significant difference for the conductance of the halitoxins-evoked current was seen but a significant depolarizing shift in the reversal potential to –20 ± 2 mV ($n = 4$), compared with the data obtained with NMDG-based solution containing 2 mM Ca^{2+} was recorded. This result suggested that the current evoked by the halitoxins involved both Ca^{2+} and NMDG influx, but that Ca^{2+} may antagonize or produce charge-screening effects to influence the flux of NMDG, analogues to mole fraction phenomena. To investigate the influx of Ca^{2+} we subsequently used fura-2 loaded DRG neurones and imaging techniques.

Investigation of Calcium Influx Evoked in DRG Neurones by Halitoxins

In this part of the study we distinguished between cultured DRG neurones and background cells such as glia and fibroblasts in the primary culture by the rise in in-

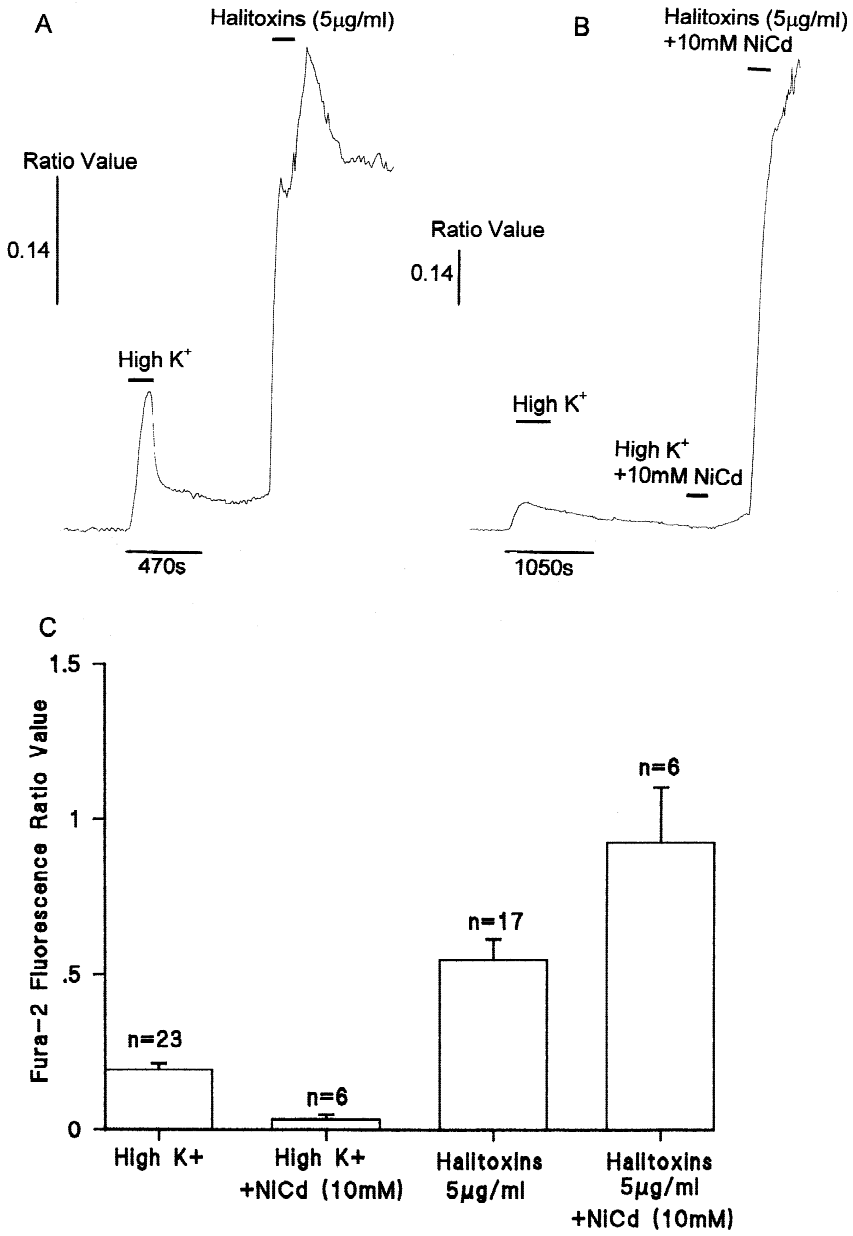


Fig. 5. Halitoxins evoked increases in intracellular Ca^{2+} in cultured DRG neurones. Increases in intracellular Ca^{2+} were detected as changes in the fluorescence ratio values using neurones loaded with the Ca^{2+}-sensitive dye, fura-2. (*A*) Mean ($n = 16$) plot for fluorescence ratio indicating increases in intracellular Ca^{2+} evoked by perfusion of NaCl-based extracellular solution containing high K^+ and halitoxins. (*B*) Mean ($n = 6$) plot for fluorescence ratio indicating increases in intracellular Ca^{2+} evoked by perfusion of NaCl-based extracellular solution containing high K^+ but no response when 10 mM $NiCl_2$ and $CdCl_2$ were simultaneously applied with high K^+. The halitoxins continued to evoke increases intracellular Ca^{2+} when they were applied simultaneously with the voltage-activated Ca^{2+} channel blockers $NiCl_2$ and $CdCl_2$. (*C*) Bar chart, showing the mean changes in fluorescence ratio values observed in response to high K^+ and halitoxins applied in the absence and presence of 10 mM $NiCl_2$ and $CdCl_2$ (NiCd).

traneuronal Ca^{2+} observed in response to high (30 mM) extracellular K^+-evoked depolarization and thus activation of voltage-gated Ca^{2+} channels. In our cultures the non-neuronal background cells load with fura-2 but do not express voltage-gated Ca^{2+} channels and thus do not show a rise in intracellular Ca^{2+} when perfused with high K^+ extracellular solution. Halitoxins (5 μg/ml) evoked substantial rises in intracellular Ca^{2+} as reflected by increases in the fura-2 fluorescence ratio values both in cultured DRG neurones (Fig. 5*A*) and in non-neuronal cells. Extracellular solution containing the voltage-gated Ca^{2+} channel inhibitors Ni^{2+} and Cd^{2+} (100 μM) failed to attenuate the K^+-evoked rise in intraneuronal Ca^{2+}. The mean values for the increases in fluorescence ratio induced by high K^+ were 0.19 ± 0.02 ($n = 23$) and 0.19 ± 0.03 ($n = 19$) under control conditions and in the presence of 100 μM Ni^{2+} and Cd^{2+}, respectively. The nature of the stimulus (high K^+) required higher concentrations of the Ca^{2+} channel inhibitors to block responses. At 10 mM Ni^{2+} and Cd^{2+} reduced the increase in fluorescence ratio produced by high K^+ to 0.03 ± 0.02 ($n = 6$), but failed to significantly attenuate the rise in intracellular Ca^{2+}-evoked by the halitoxins (Fig. 5*B* and *C*).

Investigation of Calcium Influx Evoked in F-11 Cells by Halitoxins

Undifferentiated F-11 cells were then used to further investigate the rise in intracellular Ca^{2+}-evoked by the halitoxins. In an undifferentiated state these F-11 cells were thought not to express voltage-gated Ca^{2+} channels. However, the undifferentiated F-11 cells were not a homogeneous population of cells and small but detectable

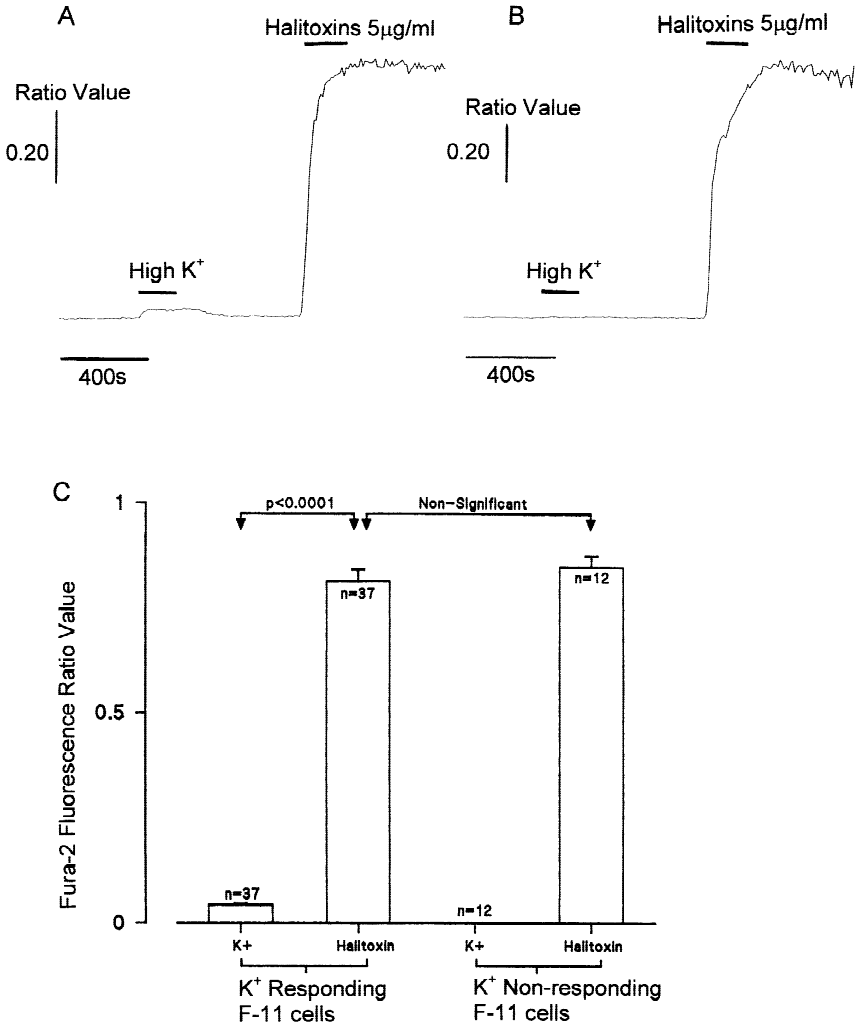


Fig. 6. Halitoxins evoked increases in intracellular Ca^{2+} in F-11 cells. (*A*) Record shows an example from an F-11 cell that showed a modest but detectable change in intracellular Ca^{2+} in response to high K^+-evoked depolarization. Subsequent application of halitoxins evoked a significantly larger rise in intracellular Ca^{2+}. (*B*) Record from an F-11 cell which did not respond to high K^+ but halitoxins still evoked a large increase in intracellular Ca^{2+}. (*C*) Bar chart, showing the mean responses to halitoxins in F-11 cells which responded to high K^+ and F-11 cells which did not respond to high K^+.

changes in intracellular Ca^{2+} in response to high K^+ were measured in 37 out of 49 cells (Fig. 6*A*). The responses to high K^+ obtained from F-11 cells were significantly smaller compared to those seen in cultured DRG neurones ($P < 0.0001$). As was seen with the primary cultures both populations of F-11 cells responded similarly to the halitoxins (5 μg/ml) with massive rises in intracellular Ca^{2+} (Fig. 6*A–C*).

Experiments were conducted to examine whether the halitoxins influenced intracellular Ca^{2+} by causing release of Ca^{2+} from intracellular stores. Surprisingly, in extracellular Ca^{2+}-free conditions (with 2.5 mM EGTA and no added $CaCl_2$) halitoxins (5 μg/ml)-evoked large transient increases in intracellular Ca^{2+} (mean change in fluorescence ratio value of 0.78 ± 0.06, $n = 21$; Fig. 7*A*). These transient Ca^{2+} responses persisted when extracellular Na^+ was replaced by NMDG to reduce the ionic disturbance produced by the halitoxins (mean change in fluorescence ratio value of 0.57 ± 0.06, $n = 13$; Fig. 7*B*). When the F-11 cells were subsequently exposed to extracellular Ca^{2+} again changes in intracellular Ca^{2+} were detected (Fig. 7*A* and *B*). This indicated that the transient nature of the responses to halitoxins observed in Ca^{2+}-free conditions was not due to a loss of fura-2 from the cells. These observations suggest that the halitoxins released Ca^{2+} from intracellular stores of F-11 cells. Experiments were then conducted with 140 mM KCl but no Ca^{2+} in the extracellular solution. Under these conditions little ionic disturbance is likely yet halitoxins evoked transient increases in intracellular Ca^{2+} with a mean change in fluorescence ratio value of 0.37 ± 0.04 ($n = 13$ out of 31 cells). However, it was clear that 18 cells did not respond to halitoxins under these conditions. Changing from the standard NaCl-based extracellular recording solution to KCl-based, Ca^{2+}-free solution evoked transient changes in intracellular Ca^{2+} in 18 out of 31 cells. Under these conditions the mean change in fluorescence ratio value was 0.24 ± 0.04 ($n = 18$). The cells that responded to perfusion with KCl-based, Ca^{2+}-free solution did not subsequently respond to halitoxins (Fig. 7*C*). One interpretation of these observations is that the KCl-based, Ca^{2+}-free solution caused depletion of intracellular Ca^{2+} stores and that these stores could not be refilled when the cells were bathed with Ca^{2+}-free extracellular solution. Hence, when the intracellular Ca^{2+} stores were depleted the halitoxins could not mo-

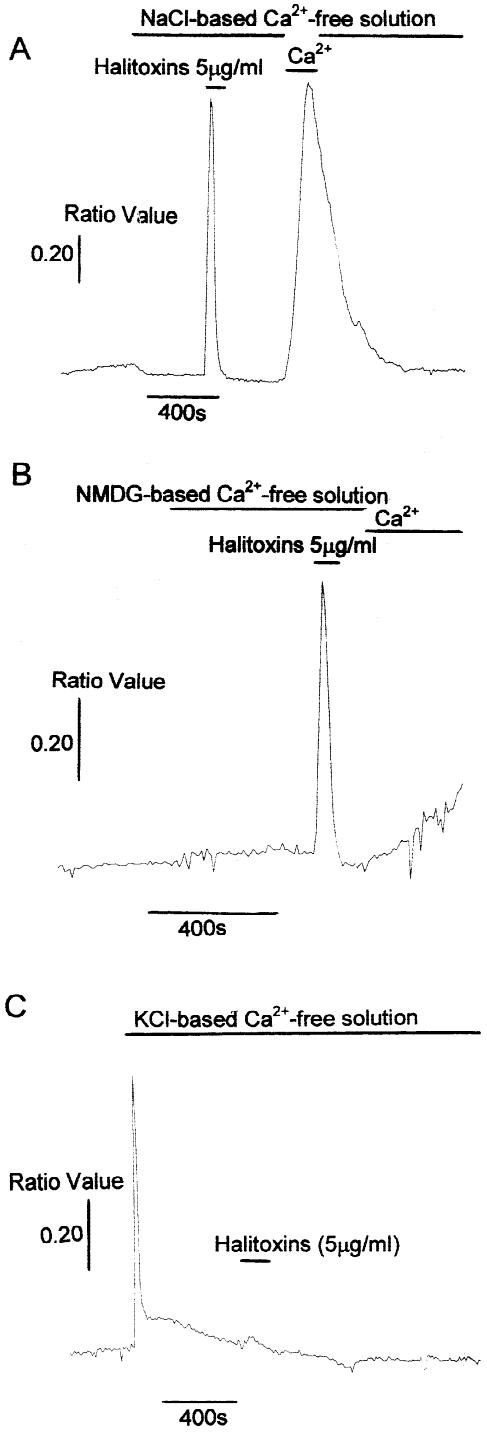


Fig. 7. Halitoxins evoked transient rises in intracellular Ca^{2+} when applied in Ca^{2+}-free extracellular solutions. (*A*) Record from an F-11 cell bathed in NaCl-based Ca^{2+}-free solution. Halitoxins evoked transient rises in intracellular Ca^{2+} and a further response was observed when NaCl-based extracellular solution containing 2 mM Ca^{2+} was perfused. (*B*) Record from an F-11 cell bathed in NMDG-based Ca^{2+}-free solution. Halitoxins evoked transient rises in intracellular Ca^{2+} and a further slow response was observed when NMDG-based extracellular solution containing 2 mM Ca^{2+} was perfused. (*C*) Example record, showing the response in an F-11 to perfusion with KCl-based, Ca^{2+}-free solution and no subsequent response to halitoxins. In all these experiments the cells were loaded with fura-2 and initially perfused with standard NaCl-based extracellular solution.

bilize intracellular Ca^{2+}. It was also noted that cells exposed to KCl-based, Ca^{2+}-free solution did not respond to 10 mM caffeine.

Pore Forming Effects of Halitoxins on Artificial Lipid Bilayers

The lack of specificity and large sizes of conductance changes observed in response to the halitoxins in this study and previous findings suggested that the 1,3 alkyl-pyridinium salts may act as pore formers rather than activators of native channels. This was investigated directly by applying the halitoxins (50 μg/ml) to artificial lipid bilayers (composed of 9:1, phosphatidylcholine and cholesterol) formed across the tip of patch pipettes. Application of the toxins, after a delay of approximately 50 sec, caused the development of a range of unitary channel-like events in lipid bilayers which had previously been "silent," or stable (Fig. 8*A*). After a variable period these events summated to give a macrocurrent. Unitary current amplitudes were measured at transmembrane potentials of −90 and +90 mV and with the estimated current reversal potential were combined to give estimates of the unitary conductances. The estimated current reversal potential was determined from the difference currents. Prior to application of halitoxins and after activation by the toxins of a stable macrocurrent, current/voltage relationships were generated between −180 mV and 0 mV, using 100 msec voltage-step commands. The control current-voltage relationship was subtracted from the relationship obtained after evoking a current with halitoxins and reversal potential of the resulting difference current was measured. The main difficulty encountered was the great variability within a single lipid bilayer between the amplitudes of the channel-like events. At −90 mV the amplitudes of the currents varied between −10 pA and −175 pA. Similarly, at +90 mV the current amplitudes varied between +15 pA and +144 pA ($n = 2$ bilayers). The estimates for the unitary conductances caused by the halitoxins were 145 pS to an upper limit of 2280 pS (Fig. 8*B* and *C*).

Discussion

Halitoxins (1,3 alkyl-pyridinium salts) produced dramatic effects on the electrophysiological properties of cultured DRG neurones, F-11 cells and artificial lipid bilayers. The actions of these toxins included irreversible depolarizations, large increases in input conductance, loss of measurable action potentials and loss of measurable voltage-activated currents. These effects are consistent with halitoxins producing ion permeable channels in cell membranes and thus large increases in input conductance on top of which it is not possible to detect

native conductances. Our results are also consistent with those from previous work showing a decline in action potential amplitude (Berlinck et al., 1996) and block of squid axon K^+ conductances (Sevcik et al., 1994) by sponge toxins. The halitoxin pores appeared to have a high permeability to divalent Ca^{2+} and the large organic monovalent cation $NMDG^+$. The Ca^{2+} permeability might have been due to depolarization-induced activation of voltage-gated Ca^{2+} channels. However, the responses to halitoxins were significantly greater than those produced by high K^+-evoked depolarizations. Additionally, the measurements from nonneuronal cells in primary cultures, DRG neurones bathed in high concentrations of Cd^{2+} and Ni^{2+} and F-11 cells which did not respond to high K^+ suggest that functional voltage-activated Ca^{2+} channels were not required for the halitoxins to produce massive influx of Ca^{2+} into cells. The Ca^{2+} component of the conductance evoked by halitoxins observed in this study also provides a potential mechanism by which toxins from *Haliclona viridis* induced concentration and Ca^{2+}-dependent release of gamma-aminobutyric acid from rat olfactory bulb neurones (Jaffe, Eisig & Sevcik, 1993).

Surprisingly, halitoxins also evoked release of Ca^{2+} from intracellular stores. This was revealed under extracellular Ca^{2+}-free conditions and may involve loss of Ca^{2+} from a variety of organelles including endoplasmic reticulum and mitochondria. The data from experiments using a high K^+ and Ca^{2+}-free extracellular solution suggested that if intracellular stores were depleted of Ca^{2+} then halitoxins failed to produce transient increases in intracellular Ca^{2+}. Several possible mechanisms may account for this effect of halitoxins on intracellular Ca^{2+} stores. Firstly, the intracellular ionic disturbance produced by the pore-forming toxins may have triggered release of Ca^{2+} from intracellular stores. However, the results obtained with NMDG- and KCl-based extracellular solutions containing no Ca^{2+} make this possibility less likely. Secondly, the organelles containing Ca^{2+} may be in very close proximity to the cell membrane so that the pore formers in spanning the cell membrane gain access to the intracellular store membranes. Release of Ca^{2+} from a store may subsequently result in Ca^{2+}-induced Ca^{2+} release and thus the substantial global responses observed. The third possibility stems from the variable size of the halitoxin molecules and that large pores formed by the larger toxin molecules may provide pathways by which smaller toxin molecules gain access to the intracellular environment and then subsequently produce channels in the intracellular Ca^{2+} stores. Previously it has been shown that serum-derived complement causes insertion of damaging membrane attack complex and increases in intracellular Ca^{2+} in oligodendroglia. This effect of serum-derived complement is dependent on the terminal complement component C9. Interestingly, intracellular Ca^{2+} oscillations in oligodendroglia are also evoked by a 26 amino acid peptide called melittin which has regions of complete homology with the C9 amino acid sequence (Wood et al., 1993). This effect may be indicative of an effect on endoplasmic reticular Ca^{2+} stores by these agents that potentially produce pore-forming lesions. Additionally, it has recently been shown that the transmembrane pore former, perforin, produces disruption of endosomal trafficking. This effect indicates that this pore former also influences the functional properties of intracellular organelles (Browne et al., 1999).

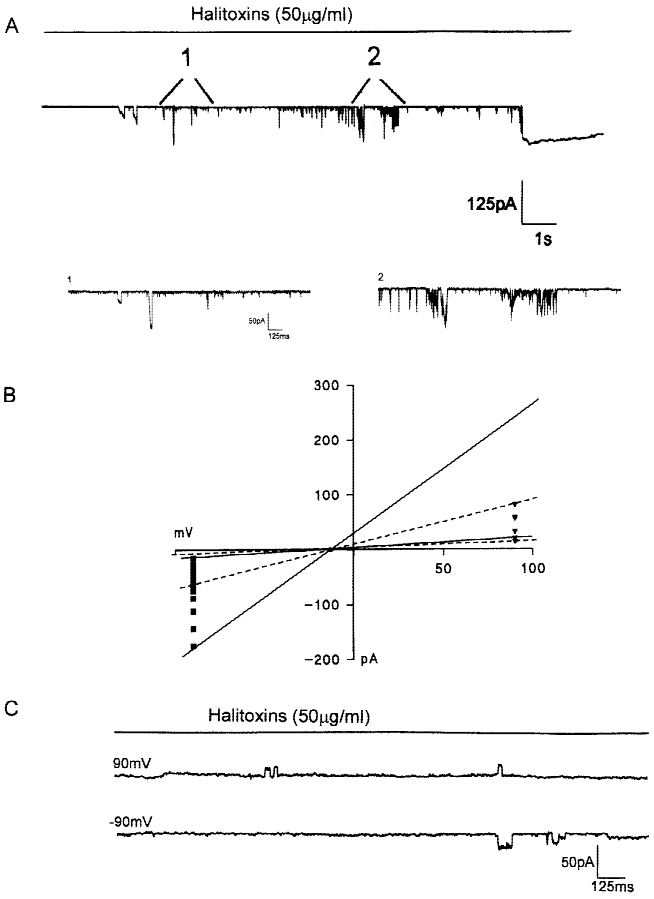


Fig. 8. Halitoxins act as pore formers on artificial lipid bilayers. (*A*) Record of channel-like activity evoked when halitoxins were applied to a lipid bilayer. The lipid bilayer was stable with no channel-like activity for 20 min before application of the halitoxins; 1 and 2 show regions of interest that have been expanded below. (*B*) Unitary current-voltage relationships from a single lipid bilayer experiment. The graph shows the great variation in channel current amplitude at –90 and +90 mV. The lines are the range of predicted slope conductances from currents measured at –90 mV (filled lines) and from currents measured at +90 mV (broken lines). (*C*) Single channel-like events produced by halitoxins when the transmembrane potential was at + 90 and –90 mV.

Direct evidence for the pore-forming properties of the halitoxins investigated in this study came from work on artificial lipid bilayers. Very varied and large unitary conductances were obtained for halitoxin channels. Some of these unitary conductance values are above the

upper theoretical limit (~300 pS) for typical native channels activated by neurotransmitters or voltage (Hille, 1992). However, large conductance channels have been reported for a number of pore-forming toxins, antibiotics and detergents. For example, α-toxin of *Staphylococcus aureus* forms two main classes of channel with conductances of 90 pS and 450 pS in 0.5 M KCl (Menestrina, 1986; Mellor, Thomas & Sansom, 1988) and an estimated pore diameter of 2–3 nm (Fussle et al., 1981; Bayley, 1997). Hemolysin from *Escherichia coli* produces channels with conductances of 500 pS in 0.15 M KCl and a pore diameter of about 1 nm (Benz et al., 1989). Ducloheir and colleagues (1998) found that antiamoebin, the fungal polypeptide antibiotic, forms pores with a range of single channel conductances (90, 480, 960, 1600 & 3200 pS, in 1 M KCl). Similarly, δ-endotoxins form channels with conductances ranging from 200 pS to about 4000 pS in 300 mM KCl (Slatin, Abrams & English, 1990). Large channels with a minimium pore diameter of 9 Å such as those formed by colicin E1 are also permeable to organic monovalent cations such as $NMDG^+$ which have dimensions of 13.3 × 5.5 Å (Bullock, Kolen & Shear, 1992). So the NMDG permeability and large conductance of the channels formed by the halitoxins is on a par with some other pore-forming agents. The variability of the channels formed may be due to the different sized 1,3 alkyl-pyridinium salts present in the preparation of halitoxins. It is not at present possible to separate out the distinct 1,3 alkyl-pyridinium salts, however they are amenable to chemical synthesis and this in the future will be an approach used to investigate the properties of the different halitoxin molecules. Further properties of these 1,3 alkyl-pyridinium salts which makes them attractive for further studies is their high water solubility and chemical stability.

From all of our experiments we predict that the halitoxin channel-like activity shows a degree of cation selectivity with K^+ and Na^+ permeabilities being similar. The permeability to Ca^{2+} appears very significant but less than for K^+ and Na^+. The large cation, $NMDG^+$ has the lowest permeability of the cations studied, never the less $NMDG^+$ conductance is clear following application of halitoxins. Detailed analysis of permeability sequences for a wide range of cations and anions is yet to be undertaken but will benefit from the preparation and use of a single and chemically defined halitoxin compound. In one important respect it is perhaps surprising that halitoxins generate cation-conducting pores given that the formal depiction of the molecules indicates polymers with strong positive charge. However, electronic calculations predict that the charge is delocalized with a small negative charge on the nitrogens of the pyridinium rings that were formally positive. Once complexed to the lipid membrane halitoxins may be nearly electrically neutral, rather than strongly positively charged. An additional consideration in the context of halitoxins forming cation permeable pores may be the interaction with counter anions and specifically Cl^- in physiological solutions and sea water.

The pore-forming effects of the halitoxin preparation was observed with quite low and biologically relevant concentrations. After partial purification, halitoxins make up at least 0.05% of the dry weight of the organism, which suggests the toxins are produced by sponges in great amounts. This could pose sponges with a particular difficulty of pore-forming toxin storage, unless the chemical defense was produced on demand. It seems likely that the halitoxins are produced in the copious enzyme-rich slime secreted by the colony after injury.

Conclusions

Our results suggest that halitoxins insert into biological and artificial lipid membranes to form ion permeable pores that allow flux of monovalent cations including the large organic ion NMDG and the divalent cation Ca^{2+}. A number of potential applications for stable pore formers are apparent (Bayley, 1997). These applications include, use as electrophysiological tools, development of novel cytotoxic molecules and possibly agents for intracellular drug delivery. The synthesis of numerous diverse and stable pore formers with distinct biophysical properties is an exciting future prospect that may be realized through the manipulation of 1,3 alkyl-pyridinium salts.

The butanolic extract of *Callyspongia ridleyi* was a generous gift from Professor Phil Crews, University of California, Santa Cruz. MALDI-TOF MS were carried out by Ian Davidson at the Proteome Unit, Institute of Medical Sciences, Aberdeen University. ESI-MS was carried out by Gary Duncan at the Rowett Research Institute, Aberdeen. RHS thanks The Wellcome Trust and Pfizer UK for support and Dr. Ian Duce for helpful discussion.

References

Albrizio, S., Ciminiello, P., Fattorusso, E., Magno, S., Pawlik, J.R. 1995. Amphitoxin, a new high molecular weight anti-feedant pyridinium salt from the Caribbean sponge *Amphimedon compressa*. *J. Nat. Prod.* **58:**647–652

Andersen, R.J., van Soest, R.W.M., Kong, F. 1996. *In:* Alkaloids: Chemical and Biological Perspectives. S.W. Pelletier, Editor. Vol. 10: pp. 430. Pergamon, Oxford

Baslow, M.H., Turlapaty, P. 1969. In vivo antitumor activity and other pharmacological properties of halitoxin obtained from the sponge *Haliclona viridis*. *Proc. West Pharmacol. Soc.* **12:**6–8

Bayley, H. 1997. Building doors into cells. *Sci. Am.* **277:**62–67

Benz, R., Schmid, A., Wagner, W., Goebel, W. 1989. Pore formation by the *Escherichia coli* hemolysin: evidence for an association-dissociation equilibrium of the pore-forming aggregates. *Infect. Immun.* **57:**887–895

Berlinck, R.G., Ogawa, C.A., Almeida, A.M., Sanchez, M.A., Malpezzi, E.L., Costa, L.V., Hajdu, E., de Freitas, J.C. 1996. Chemical and pharmacological characterization of halitoxin from *Amphimedon viridis* (Porifera) from the southeastern Brazilian coast. *Comp. Biochem. Physiol.* **115C:**155–163

Browne, K.A., Blink, E., Sutton, V.R., Froelich, C.J., Jans, D.A., Trapani, J.A. 1999. Cytosolic delivery of granzyme B by bacterial toxins: evidence that endosomal disruption, in addition to transmembrane pore formation, is an important function of perforin. *Mol. Cell Biol.* **19:**8604–8615

Bullock, J.O., Kolen, E.R., Shear, J.L. 1992. Ion selectivity of colicin E1: II. Permeability of organic cations. *J. Membrane Biol.* **128:** 1–16

Coronado, R., Latorre, R. 1983. Phospholipid bilayers made from monolayers on patch-clamp pipettes. *Biophys. J.* **43:**231–236

Crews, P., Cheng, X.C., Adamczeski, M., Rodriguez, J., Jaspars, M., Schmitz, F.J., Traeger, S.C., Pordesimo, E.O. 1994. 1,2,3,4-Tetrahydro-2-N-Methyl-8-Hydroxymanzamine A: an Alkaloid Salt from the Sponge *Petrosia contignata*. *Tetrahedron* **50:**13567–13574

Davies-Coleman, M.T., Faulkner, D.J., Dubowchik, G.M., Roth, G.P., Polson, C., Fairchild, C. 1993. A new EGF-active polymeric pyridinium alkaloid from the sponge *Callyspongia fibrosa*. *J. Org. Chem.* **58:**5925–5930

Duclohier, H., Snook, C.F., Wallace, B.A. 1998. Antiamoebin can function as a carrier or as a pore-forming peptaibol. *Biochim. Biophys. Acta* **1415:**255–260

Fussle, R., Bhakdi, S., Sziegoleit, A., Tranum-Jensen, J., Kranz, T., Wellensiek, H.J. 1981. On the mechanism of membrane damage by *Staphylococcus aureus* alpha-toxin. *J. Cell Biol.* **91:**83–94

Hamill, O.P., Marty, A., Neher, E., Sakmann, B., Sigworth, F. 1981. An improved patch clamp technique for high resolution current recordings from cells and cell free membrane patches. *Pfluegers Arch.* **391:**85–100

Hille, B. 1992. Ionic Channels of Excitable Membranes. 2nd Edition, pp. 332. Sinauer Associates, Sunderland, MA

Jaffe, E., Eisig, M., Sevcik, C. 1993. Effect of a toxin isolated from the sponge *Haliclona viridis* on the release of gamma-aminobutyric acid from rat olfactory bulb. *Toxicon* **31:**385–396

Jaspars, M., Pasupathy, V., Crews, P. 1994. A tetracyclic diamine alkaloid, halicyclamine A, isolated from the marine sponge *Haliclona sp. J. Org. Chem.* **59:**3253–3255

Jaspars, M., Rali, T., Laney, M., Schatzman, R.C., Diaz, M.C., Schmitz, F.J., Pordesimo, E.O., Crews, P. 1994. The search for inosine 5′ dehydrogenase (IMPDH) inhibitors from marine sponges. Evaluation of the bastadin alkaloids. *Tetrahedron* **50:**7367–7374

Mellor, I.R., Thomas, D.H., Sansom, M.S. 1988. Properties of ion channels formed by *Staphylococcus aureus* delta-toxin. *Biochim. Biophys. Acta* **942:**280–294

Menestrina, G. 1986. Ionic channels formed by *Staphylococcus aureus* alpha-toxin: voltage-dependent inhibition by divalent and trivalent cations. *J. Membrane Biol.* **90:**177–190

Platika, D., Boulos, M.H., Baizer, L., Fishman, M.C. 1985. Neuronal traits of clonal cell lines derived by fusion of dorsal root ganglia neurons with neuroblastoma cells. *Proc. Natl. Acad. Sci. U.S.A.* **82:**3499–3503

Schmidt, M.W., Baldridge, K.K., Boatz, J.A., Elbert, S.T., Gordon, M.S., Jensen, J.J., Koseki, S., Matsunaga, N., Nguyen, K.A., Su, S., Windus, T.L., Dupuis, M., Montgomery, J.A. 1993. General atomic and molecular electronic-structure system. *J. Comput. Chem.* **14:**1347–1363

Schmitz, F.J., Hollenbeak, H.H., Campbell, D.C. 1978. Marine natural products: halitoxin, toxic complex of several marine sponges of the genus *Haliclona*. *J. Org. Chem.* **43:**3916–3922

Sepcic, K., Guella, G., Mancini, I., Pietra, F., Dalla Serra, M., Menestrina, G., Tubbs, K., Macek, P., Turk, T. 1997. Characterization of anticholinesterase-active 3-alkylpyridinium polymers from the marine sponge *Reniera sarai* in aqueous solutions. *J. Nat. Prod.* **60:**991–996

Sevcik, C., Alvarez-Vasquez, F., Saavedra, J.A., Cordovez, G. 1986. Blockage of resting potassium conductance in frog muscle fibers by a toxin isolated from the sponge *Haliclona viridis*. *Toxicon* **24:**851–860

Sevcik, C., Garcia-Rodriguez, A.I., D'Suze, G., Mijares, A.J. 1994. Specific blockage of squid axon resting potassium permeability by *Haliclona viridis* (*Porifera: Haliclonidae*) toxin (HvTX). *Toxicon* **32:**773–788

Slatin, S.L., Abrams, C.K., English, L. 1990. Delta-endotoxins form cation-selective channels in planar lipid bilayers. *Biochem. Biophys. Res. Commun.* **169:**765–772

Wood, A., Wing, M.G., Benham, C.D., Compston, D.A. 1993. Specific induction of intracellular calcium oscillations by complement membrane attack on oligodendroglia. *J. Neurosci.* **13:**3319–3332

Pergamon

Tetrahedron 57 (2001) 3199–3207

TETRAHEDRON

Metal binding of *Lissoclinum patella* metabolites. Part 2: Lissoclinamides 9 and 10

Linda A. Morris,[a] Bruce F. Milne,[a] Marcel Jaspars,[a,*] J. Jantina Kettenes-van den Bosch,[b] Kees Versluis,[b] Albert J. R. Heck,[b] Sharon M. Kelly[c] and Nicholas C. Price[c]

[a]*Department of Chemistry, Marine Natural Products Laboratory, University of Aberdeen, Old Aberdeen, AB24 3UE Scotland, UK.*
[b]*Department of Biomolecular Mass Spectrometry, Utrecht Institute for Pharmaceutical Sciences (UIPS) and Bijvoet Center for Biomolecular Research, Utrecht University, Sorbonnelaan 16, 3584 CA Utrecht, The Netherlands.*
[c]*IBLS Division of Biochemistry and Molecular Biology, Joseph Black Building, University of Glasgow, Glasgow, G12 8QQ Scotland, UK.*

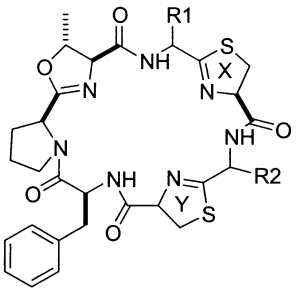
Received 1 December 2000; revised 15 January 2001; accepted 1 February 2001

Abstract—Studies on the Thz, Thn and Oxn containing cyclic peptides, lissoclinamides 9 (**9**) and 10 (**10**) isolated from the Indo-Pacific ascidian (seasquirt) *Lissoclinum patella* have delineated their metal binding selectivity. MS and CD competition studies show that lissoclinamide 10 (**10**) shows selectivity for Cu^{2+} in the presence of an excess of Zn^{2+} whereas lissoclinamide 9 (**9**) is less selective for Cu^{2+}. Comparison of the solution state conformations derived from nOe restrained molecular dynamics and additional Monte–Carlo conformational searches suggested binding environments for the Cu^{2+} which confirmed the MS measurements and suggested a reason for the selectivity in the case of lissoclinamides 9 and 10. © 2001 Elsevier Science Ltd. All rights reserved.

1. Introduction

In this paper we shall be concerned with the *Lissoclinum patella* metabolites lissoclinamide 9 and 10, members of the lissoclinamide group of compounds (**1**–**11**), composed of seven amino acid residues containing thiazole (Thz), thiazoline (Thn) and oxazoline (Oxn) ring systems. The isolation and structure determination of lissoclinamides 9 (**9**) and 10 (**10**) was the subject of a previous publication.[1] Several natural and synthetic lissoclinamides were found to be biologically active. Wipf and co-workers showed that the replacement of Thn rings with Oxn rings decreased activity to a greater extent than replacement of Oxn rings with Thn rings.[2] This study further showed that it was not only the individual components of the macrocycle that conferred high activity but rather the overall conformation of the molecules. The structure activity relationship is also demonstrated when comparing lissoclinamides 4 (**4**) and 5 (**5**). These compounds differ only in the oxidation state of a single Thz unit but this difference makes lissoclinamide 5 (**5**) two orders of magnitude less cytotoxic than lissoclinamide 4 (**4**) against bladder carcinoma (T24) cells.[3]

The lissoclinamides resemble 21-azacrown-7 macrocycles but no investigations into their metal binding properties have been reported, unlike for the patellamides for which extensive binding studies have been published in the literature (see preceding paper).



1 Lissoclinamide 1 R1=L-Val R2=D-Ile X=Y=thiazole
2 Lissoclinamide 2 R1=D-Ile R2=D-Ala X=thiazolineY=thiazole
3 Lissoclinamide 3 R1=D-Ile R2=L-Ala X=thiazolineY=thiazole
4 Lissoclinamide 4 R1=L-Val R2=D-Phe X=thiazolineY=thiazole
5 Lissoclinamide 5 R1=L-Val R2=D-Phe X=Y=thiazole
6 Lissoclinamide 6 R1=D-Val R2=D-Phe X=thiazolineY=thiazole
7 Lissoclinamide 7 R1=D-Val R2=D-Phe X=Y=thiazoline
8 Lissoclinamide 8 R1=Val R2=Phe X=thiazolineY=thiazole
11 Ulicyclamide R1=L-Ile R2=D-Ala X=Y=thiazole

Keywords: complexation; natural products; circular dichroism; mass spectrometry.

* Corresponding author. Tel.: +44-1224-272-895; fax: +44-1224-272-921; e-mail: m.jaspars@abdn.ac.uk

0040–4020/01/$ - see front matter
PII: S0040-4020(01)00178-8

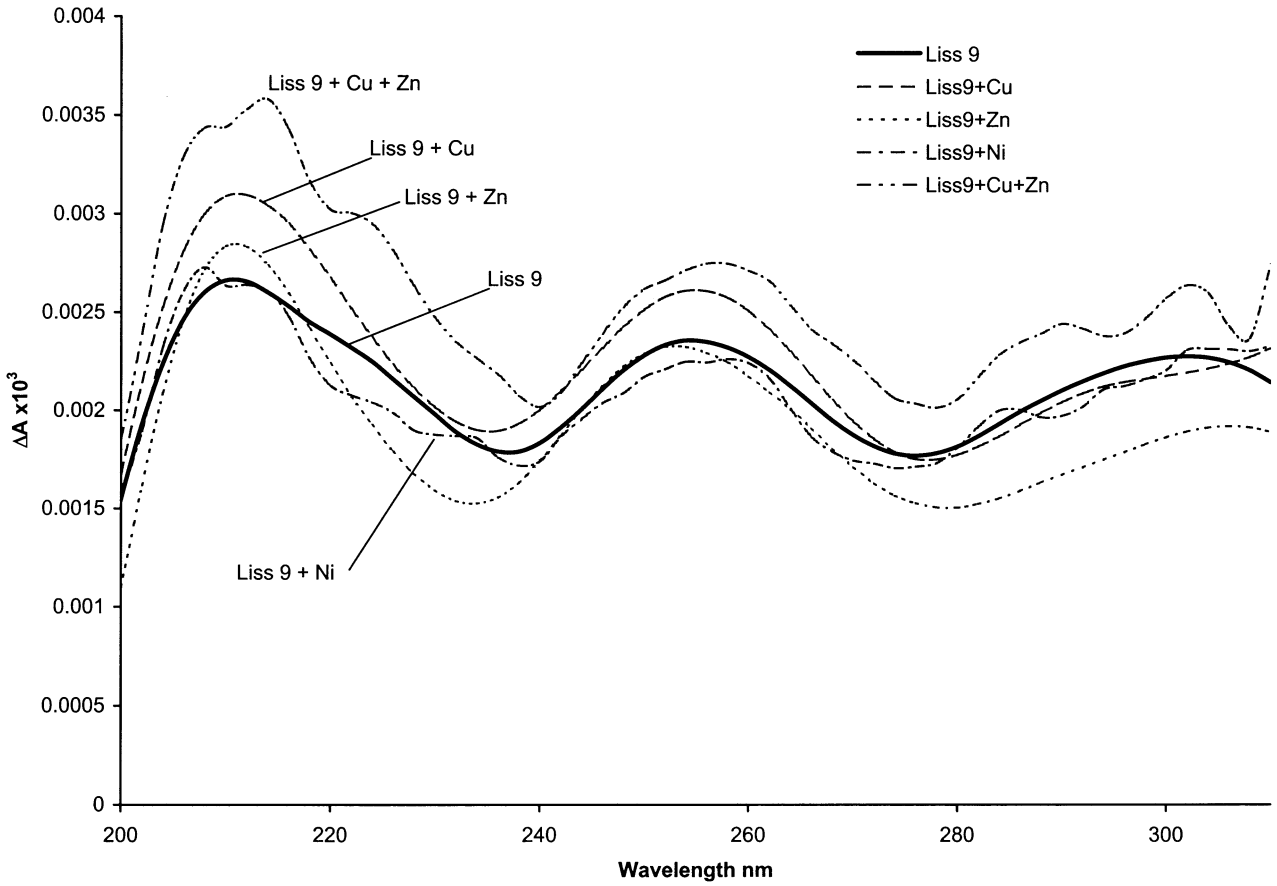
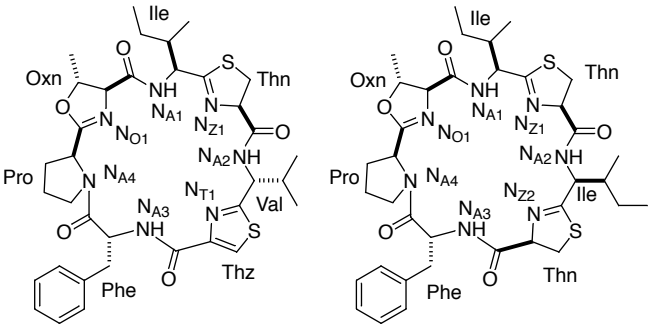


Figure 1. CD spectra of **9** at 0.002 mg/mL in MeOH plus 2 equiv. of copper, zinc, nickel and a mixture of copper and zinc.



9 **10**

2. Results and discussion

2.1. Circular dichroism studies

The first CD spectra obtained were those of lissoclinamides 9 and 10 (**9**, **10**) in methanolic solution (Figs. 1 and 2, bold traces). The spectra are of much lower intensity than the patellamide spectra and there are no prominent maxima or minima (see preceding paper). Qualitative comparisons to standard protein CD spectra[4] indicate that a weak type II β-turn is present in lissoclinamide 9 (λ_{max}=209, 254 nm) similar to that found in the 'figure of eight' conformations of the patellamides.[5] This β-turn is induced by the Oxn moiety as shown by Wipf et al.[2] in studies on lissoclinamide 7 (**7**) (Table 1). A Monte–Carlo conformational search on lissoclinamide 9 indicates that in the lowest energy form (**9** MC_L, Fig. 3) a hydrogen bond is present between Phe C=O and Ile N–H (N_{A1}–H) and that the ϕ and ψ angles fall in the acceptable range for type II β -turns (Table 1).[6] However in an nOe restrained molecular dynamics structure of lissoclinamide 9 this β-turn is absent (Table 1).[1] Re-examination of the Monte–Carlo molecular modeling data shows that there is a family of structures (**9** MC_H), of which the energy is 8 kJ/mol higher than the global minimum and which shows good overlap with the nOe derived structure in the Pro-Oxn part of the molecule (full backbone RMSD= 0.92 Å; RMSD for Pro–N–Ile–N backbone=0.11 Å). The absence of a β-turn in thermally accessible conformations explains why the intensity of the β-turn signature in the CD spectrum is so low.

The CD spectrum of lissoclinamide 10 shows features consistent with a random coil type structure (Fig. 2, bold trace).[4] The greater flexibility of this molecule, due to the presence of two Thn moieties, is evident in the Monte–Carlo conformational search. The lowest energy structure (**10** MC_L, Fig. 4) does not contain the Phe C=O to Ile N–H (N_{A1}–H) hydrogen bond necessary for a β-turn whereas it is present in a structural family with a slightly higher energy (**10** MC_H, ~1 kJ/mol) (Table 1). Examination of families of conformations within a thermally accessible temperature range shows that tight hydrogen bonded turns can also be formed under the influence of each Thn moiety, although

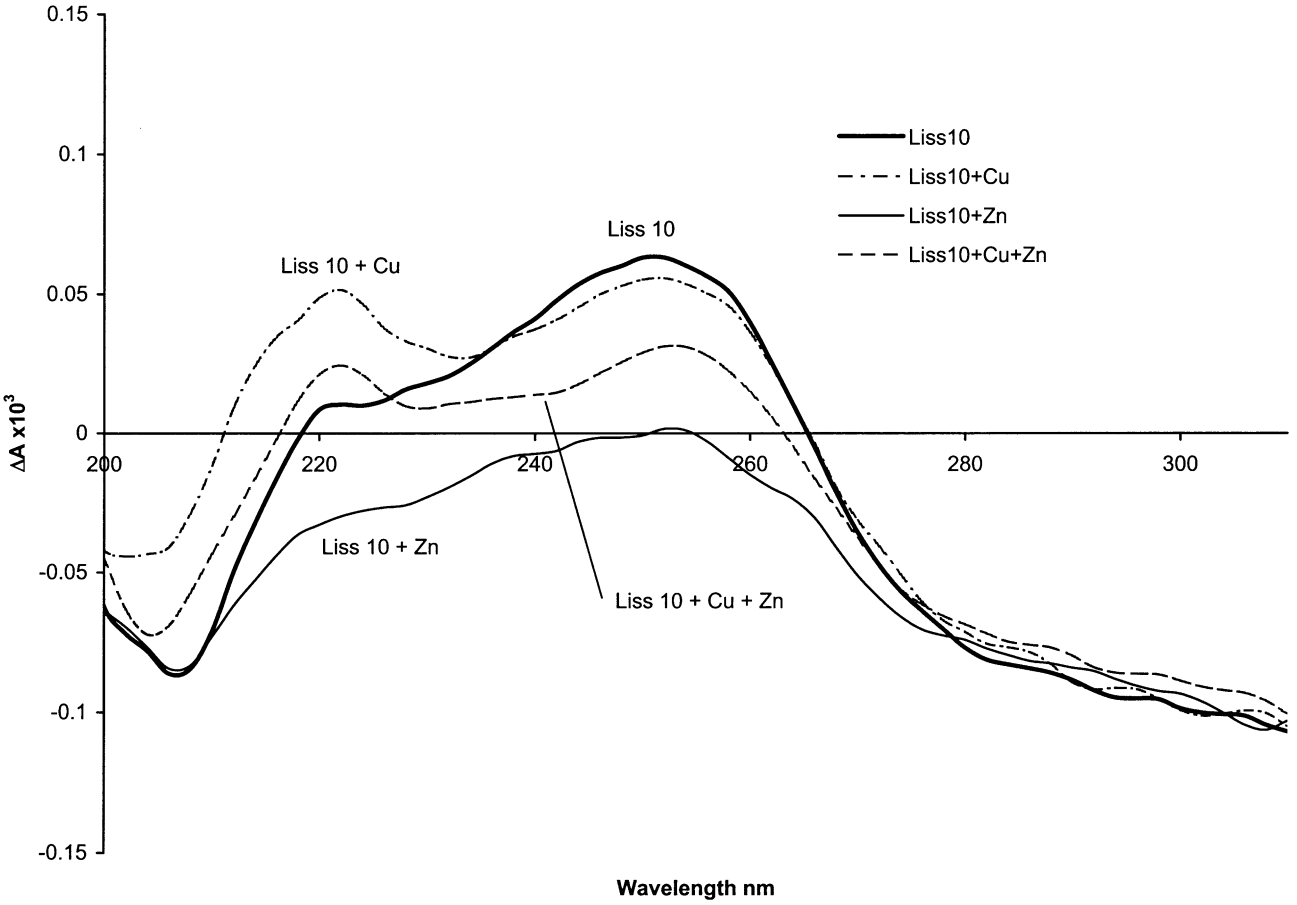


Figure 2. CD spectra of **10** at 0.02 mg/mL in MeOH plus 2 equiv. of copper, zinc and a mixture of copper and zinc.

they do not fit the standard types of turn (e.g. H-bond IleNH (N_{A2}–H) –OxnC═O; $\phi_{Ile}=-30°$; $\psi_{Ile}=109°$; $\phi_{Thn}=137°$; $\psi_{Thn}=-40°$).[6] The absence of a dominant low energy conformation therefore explains why the CD spectrum indicates the presence of a random coil.

Addition of 2 equiv. of $CuCl_2$, $ZnCl_2$ or $NiCl_2$ to a solution of lissoclinamide 9 (**9**) caused little observable change in the CD maxima at 209 and 254 nm indicating that little or no conformational change was taking place (Fig. 1). A small increase in ΔA at λ_{max} 209 and 254 was noted on addition of copper, but the spectral traces for **9**/Zn^{2+} and **9**/Ni^{2+} essentially followed that of native **9**. It was therefore not possible to discern if **9** was binding to Zn^{2+} or Ni^{2+}. The invariance of the CD spectra suggested that the conformation of **9** was very stable. The small change at 209 and 254 nm in a titration of **9** with Cu^{2+}, monitored by CD, was used to estimate binding constants. It was found that the first binding constant $K_1=1.4\pm0.2\times10^4$, similar to that for patellamide C (see preceding paper) and the second binding constant was much weaker, $K_2=125\pm4$.

Addition of 2 equiv. of $ZnCl_2$ to a solution of lissoclinamide 10 (**10**) decreased the intensity of the CD spectrum but did not change its shape relative to that of native **10**, suggesting again that little change of conformation takes place (Fig. 2). However, addition of 2 equiv. of $CuCl_2$ to a solution of **10** increased ΔA at the 220 nm maximum whilst leaving the 255 nm maximum unaffected. This indicates that the addition of Cu^{2+} alters the conformation of lissoclinamide **10**. Lack of material effectively ruled out estimating binding constants for **10** through a CD monitored titration.

Behaviour similar to that of patellamide A (see preceding paper) was observed on adding a mixture of Cu^{2+} and Zn^{2+} to **9** (Fig. 1). Only a very small change was observed, and as

Table 1. Type II β turns in lissoclinamides

	ϕ_{Pro} (°)	ψ_{Pro} (°)	ϕ_{Oxn} (°)	ψ_{Oxn} (°)	H-bond (Å)	E_{MM} (kJ/mol)
Ideal	−60.0	120.0	80.0	0.0	1.8–2.5	
Lissoclinamide 9 NOE	−79.2	10.0	110.8	16.6	4.5	
Lissoclinamide 9 MC_L	−72.8	104.5	97.2	−3.7	2.0	156.5
Lissoclinamide 9 MC_H	−76.5	11.0	117.5	18.6	4.8	164.2
Lissoclinamide 10 MC_L	−73.8	104.2	97.0	42.0	2.8	155.7
Lissoclinamide 10 MC_H	−71.6	112.2	99.5	0.9	2.1	156.8
Lissoclinamide 7 XRD	−59.3	127.0	110.9	−24.7	2.2	

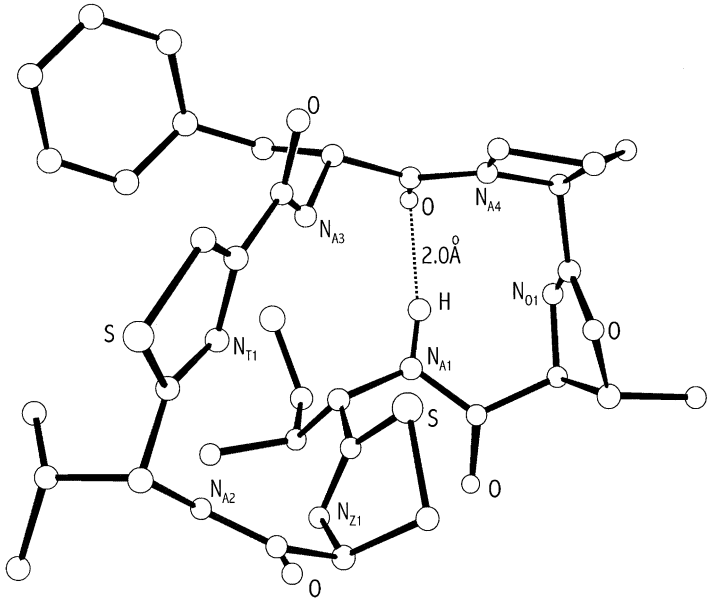


Figure 3. Global minimum **9** MC_L from molecular mechanics Monte–Carlo conformational analysis of **9** showing β-turn H-bond.

the CD spectra of **9**, **9**/Cu^{2+} and **9**/Zn^{2+} were similar in appearance no firm conclusions could be drawn. Addition of a mixture of Cu^{2+} and Zn^{2+} to lissoclinamide 10 (**10**) did indicate that some preferential binding to Cu^{2+} was taking place (Fig. 2). The shape of the resulting CD curve was very similar to that of **10**/Cu^{2+}, suggesting that **10** preferentially binds to Cu^{2+} in the presence of Zn^{2+}, a fact which was later confirmed by mass spectrometry.

3. Mass spectrometry

As **9** and **10** had been found to bind both copper and zinc in the circular dichroism study, it was decided to further investigate these complexes by mass spectrometry in order to ascertain their exact nature and which, if any, counter-ions were incorporated into the complexes. Accurate mass measurements and analysis of isotope ratio patterns were used to assign the components of a complex. We first obtained accurate mass and fragmentation data for uncomplexed **9** and **10**.[1] The copper and zinc complexes of each were then studied in detail, including their fragmentation patterns. In addition, MS was used to monitor some competition experiments.

3.1. Mass spectrometry of Cu and Zn complexes

The main species occurring in the ion trap mass spectra of lissoclinamide 9 and 10 with 2 equiv. of Cu^{2+} or Zn^{2+} are presented in Table 2. The mass spectra of lissoclinamide solutions containing Cu are simple, with few species

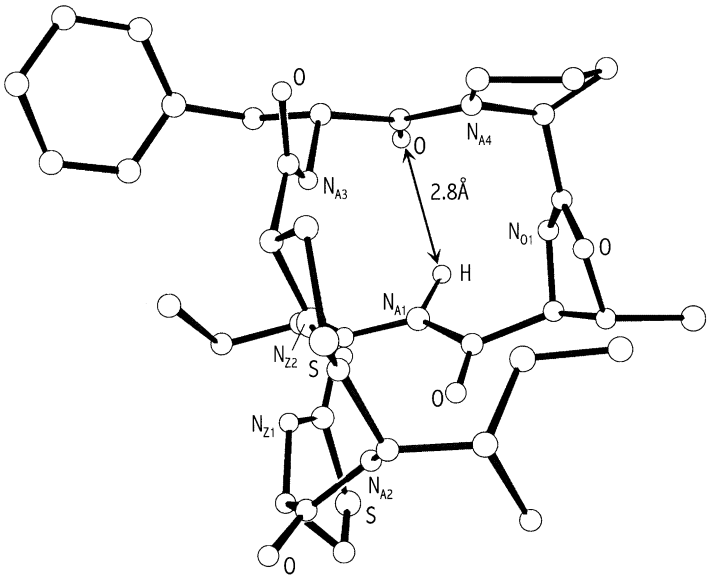


Figure 4. Global minimum **10** MC_L from molecular mechanics Monte–Carlo conformational analysis of **10** in the same orientation as **9** in Fig. 3 showing the increased distance between N_{A1}–*H* and Phe–C═*O*.

Table 2. Main species detected in the MS of Lissoclinamide 9 and Lissoclinamide 10 plus 2 equiv. Cu^{2+} or 2 equiv. Zn^{2+} solution. The relative intensities are from selective ion monitoring traces. Base peaks (100%) are in **bold** type

	m/z, Δ (%)[a]			
Species	**9**+Cu	**9**+Zn	**10**+Cu	**10**+Zn
Singly charged species				
$[M+H]^+$	708.3001	708.3001	724.3341	724.3341
	2.7 (40)	2.7 (35)	1.7 (50)	1.7 (25)
$[M+H+H_2O]^+$		**726**		
Cu containing species				
$[M-H+Cu]^+$	769.2132		**785.2444**	
	2.7 (90)		0.0	
$[M-H+Cu+H_2O]$	787 (75)		803 (10)	
$[M-H+Cu+MeOH]^+$	801 (65)			
$[M+CuCl]^+$	805 (45)			
$[M-2H+Cu_2Cl]^+$	866 (55)			
$[M-2H+Cu_2Cl+H_2O]^+$	884 (75)			
$[M-2H+Cu_2Cl+CO_2]^+$			926[b,c] (50)	
Zn containing species				
$[M-H+Zn]^+$		770 (25)		786 (15)
$[M+ZnCl]^+$		806 (25)		823 (10)
Doubly charged species				
$[M-2H+2Cu]^{2+}$	415.5 (60)			
$[M+Zn]^{2+}$		385.7 (35)		**393.5**

[a] Δ in milli-amu from calculated mass, relative intensity (%); accurate masses were measured with a quadrupole/time-of-flight instrument; all other masses with an ion trap instrument in selective ion monitoring mode. Only relative intensities >10% are included, except when accurate masses were measured.
[b] Determined by isotope ratio model for $C_{37}H_{47}N_7O_7S_2Cu_2Cl$ exp: 926 (90); 927 (50); 928 (100); 929 (50); 930 (50); 931 (25); 932 (12). calc: 926 (72); 927 (35); 928 (100); 929 (42); 930 (51); 931 (21); 932 (10).
[c] Present in a sample containing both Cu^{2+} and Zn^{2+}.

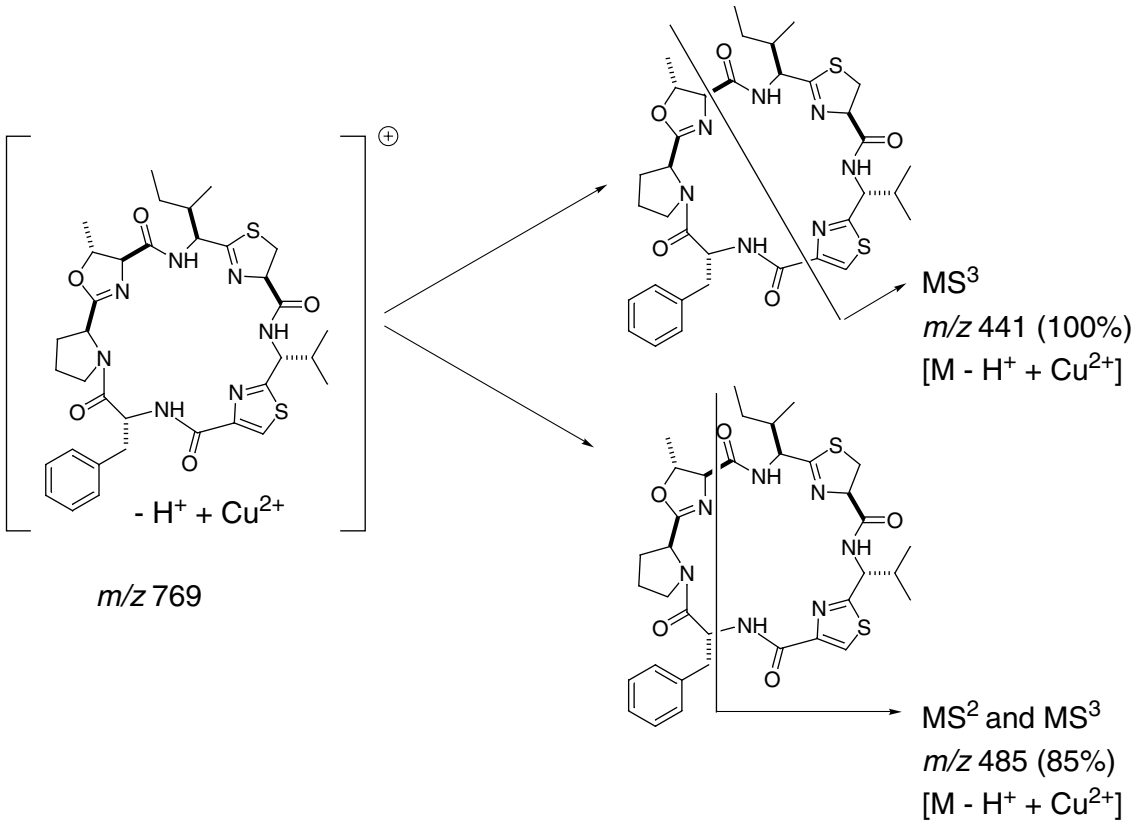


Figure 5. MS^n fragmentations of $[\mathbf{9}\text{-}H+Cu]^+$ which led to the identification of the Cu^{2+} binding site in **9**.

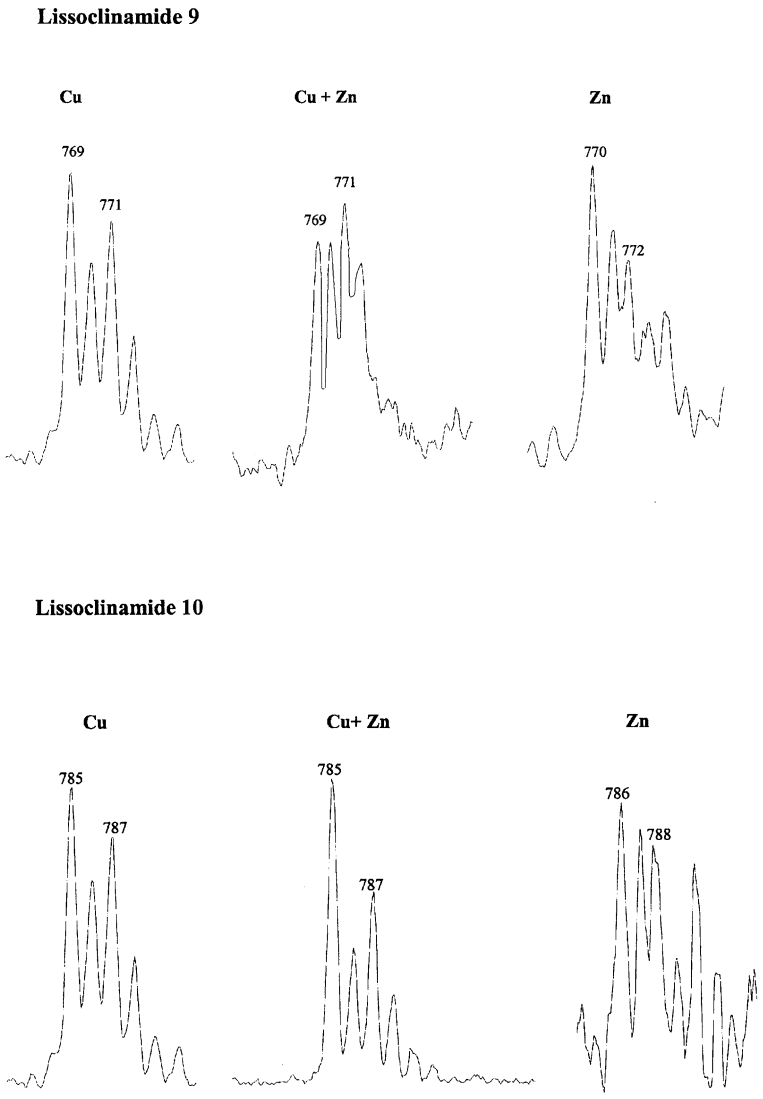


Figure 6. Selected ion monitoring spectra of $[M-H+X]^+$ (M=lissoclinamide; X=metal ion) after addition of Cu^{2+}, $Cu^{2+}+Zn^{2+}$ or Zn^{2+}.

present, whereas those of Cu containing lissoclinamide 9 and in particular those of Zn containing samples of both **9** and **10** are complex. In contrast to the mass spectra of the Cu complexes, those of the Zn complexes generally showed more multiply charged ions. Obviously, Zn has a higher preference for binding without the abstraction of hydrogen. The same was observed for patellamides A and C (see preceding paper).

For lissoclinamide 9 (**9**/Cu^{2+}), water, methanol and Cl^- adducts are far more pronounced than for the patellamides (see preceding paper). A few doubly charged ions were present as well, the strongest being at *m/z* 415.5 for $[Liss9\text{-}2H+Cu_2]^{2+}$. The MS of **10**/$Cu^{2+}$ shows $[Liss10\text{-}H+Cu]^+$ and its water adduct, as well as a dicopper complex, in combination with a considerable amount of uncomplexed **10**.

MS^n data of native **9** and **10** have been previously reported.[1] In MS^2 of **9** the main losses (apart from some 18 and 28) are 147 (PheCHNHCO, base peak), 69 (Pro) and 28 (CO). MS^2 of the Cu complex of **9** (*m/z* 769) shows loss of 91 ($PheCH_2$) as the base peak. All other fragmentations, apart from loss of 27/28 (CHN, CO) involve Pro, Oxn and $PheCH_2CH$, but the amide group ($N_{A3}H$–C═O) between Phe and Thz appears to remain intact. MS^3 of *m/z* 769 shows two large Cu containing fragments (Fig. 5). This suggests that Cu binds to the part of the molecule containing Ile, Thn, Val, and Thz. In MS^2, the Cu_2Cl complex (*m/z* 866) loses both HCl and CuCl, but the second Cu remains. These results indicate specific binding of one Cu^{2+} by **9** with a second more loosely bound, in agreement with the CD measurements of copper binding by this peptide.

In the MS^2 of native **10**, fragmentation is more pronounced. The base peak occurs for the loss of Pro+CO, but the relative intensity of signals corresponding to, amongst others, loss of 18, 28, 44, 69 (Pro), 85 (Thn) and 147 (PheCHNHCO) are over 50%. In the MS^2 of the Cu complex (*m/z* 785) the main fragmentation is loss of 83 (Oxn) and/or 71 (Pro+2H). Also, some side chain fragmentation (loss of 29, 57 and 91) is observed, which is absent in the MS^2 of **9**. However, the relative intensity of the corresponding fragment ions is below 20%. The fragmentation for the Cu complexes of both **9** and **10** strongly suggest that the binding site for copper is N_{Z1}, N_{A2} and N_{T1} for **9** and N_{Z1},

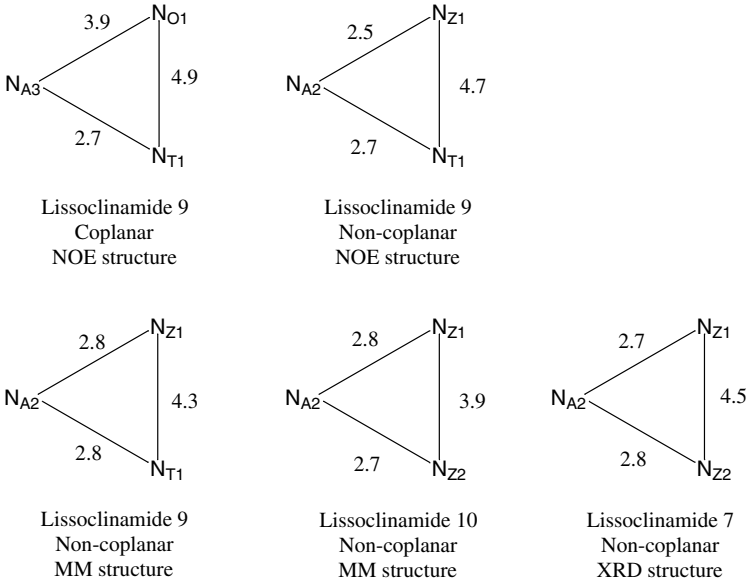


Figure 7. Distances in Å between and relative orientation of nitrogen atoms available for copper binding in **9**, **10**.

N_{A2} and N_{Z2} for **10**. That only one amide nitrogen is involved in binding is attested to by the fact that only one amide hydrogen is abstracted in the binding process, thus ruling out the other possibilities from the MS data for **9** (Fig. 5) of N_{A1}–N_{Z1}–N_{A2} and N_{A2}–N_{T1}–N_{A3}.

The mass spectrum obtained for a mixture of **9** and Zn^{2+} showed little evidence for binding of zinc. The major peaks are presented in Table 2. The water adduct of **9** was of much greater intensity than that seen in the spectrum of the native peptide, but this may be due to water being present in the $ZnCl_2$ used. The Zn complexes tended to fragment in a way similar to the native peptide, losing CO and other ring fragments, such as Thz and Pro.[1]

In the mass spectrum of **10** with Zn^{2+} only minor peaks for singly charged species such as *m/z* 786 $[Liss10\text{-}H+Zn]^+$ and its HCl adduct are observed, however, the peaks for doubly charged species for example that at *m/z* 393.5 $[Liss10+Zn]^{2+}$ are more intense, and there was also a significant amount of unbound peptide present. When fragmented, the $[Liss10+Zn]^{2+}$ species loss part of the Ile sidechain and part of the ring but retained its dipositive charge. As with patellamide C/Cu^{2+} (see preceding paper), **10**/Zn^{2+} showed no loss of the metal on fragmentation. Lissoclinamide **10** seems to bind smaller amounts of zinc compared to copper but in a more specific manner than lissoclinamide **9**.

Addition of both Cu^{2+} and Zn^{2+} to **9** and **10** seems to show mixtures of Zn^{2+} and Cu^{2+} complexes for **9**, with several multiply charged species remaining, but for **10** the mono Cu^{2+} complex at *m/z*785 is the main species (Fig. 6). Interestingly, the spectrum of **10** with Cu^{2+} and Zn^{2+} also shows a peak at *m/z* 926, which could be the CO_2 adduct of **10**+Cu_2Cl (Table 2), similar to that observed for ascidiacyclamide.[7]

4. Structural studies

Ab initio molecular mechanics conformational analyses were performed for **9** and **10** in order to assess the geometry of potential metal-binding sites within the molecule. These sites were thought to involve the Thz/Thn and intervening amide nitrogens. Global minima were located for lissoclinamides 9 and 10 with E_{MM2^*}=156.53 and 155.72 kJ/mol, respectively. For **9** the 19 lowest energy structures fell within an RMSD=0.25 Å (Fig. 3), and for **10**, the 99 lowest energy structures fell within an RMSD of 0.1 Å (Fig. 4), indicating that global minima had been located.

In the nOe restrained dynamics structure of **9**[1], a coplanar Cu^{2+} binding motif made up of N_{T1}, N_{A3} and N_{O1} (cf. the 'TAO' motif in the patellamide, see preceding paper) can be observed but the distances for binding to Cu^{2+} are non-ideal (Fig. 7). MS studies already indicated that N_{O1} was probably not involved in Cu^{2+} binding. MS fragmentation studies, allowed the identification of Cu^{2+} binding sites in **9** and **10**, and these were then studied in the global energy minima obtained from the Monte–Carlo conformational searches. The MS of **9** suggested that a binding site might present itself between the Thz N_{T1}, intervening amide N_{A2} and Thn N_{Z1} nitrogens, the 'TAZ' motif. The distances measured from the nOe and conformational search structures were similar to those in the TAO motif in native ascidiacyclamide (see preceding paper). In **10**, a similar motif, comprising of two Thn nitrogens N_{Z1} and N_{Z2} and the intervening amide nitrogen N_{A2} could be identified (the 'ZAZ' motif) with distances and geometry similar to the

dicopper ascidiacyclamide (see preceding paper). The geometries of the TAZ motif in **9** and the ZAZ motif in **10** were compared to the ascidiacyclamide TAO motif and in **10** was found to be closer to the ideal geometry for binding to Cu^{2+}. The two Thn nitrogens N_{Z1} and N_{Z2} are pointing directly at each other and the centre of the potential binding site. In **9**, the geometry was less optimal and the Thz N_{T1} and Thn N_{Z1} could be seen to be less well directed towards the centre of the binding site. In both the molecules, the binding amide nitrogen N_{A2} was easily accessible from outside the macrocycle. In addition, the presence of two Thn moieties in **10**, increases the flexibility around ZAZ motif and allows it to adopt the ideal binding geometry for Cu^{2+}. In **9**, the substitution of one Thn for a Thz increases rigidity around the TAZ motif, thus not allowing it to adopt an ideal Cu^{2+} binding geometry. At the moment, the identity of the second binding Cu^{2+} site in **9** and **10** is still ambiguous.

In **9**, the β-turn hydrogen bond (Fig. 3) does not involve the amide N_{A2}–*H* of the TAZ motif. In **10** (MC_H) however, a hydrogen bond could be seen to involve the amide N_{A2}–*H* that would have to be removed for full occupation of the ZAZ motif by Cu^{2+}. This hydrogen bond would have the effect of elongating and weakening the N–H bond, therefore making deprotonation easier. This, combined with the Thn geometry would suggest that **10** would be a better ligand for copper than **9** in that less energy would be required for complexation of the metal. The geometry of the Cu^{2+} ZAZ binding site in lissoclinamide 7 (**7**), taken from the crystal structure,[2] is found to be similar to that of **9** (Fig. 7).

In the lissoclinamides it is probable that the inclusion of two thiazoline moieties is necessary for efficient and strong binding to two Cu^{2+}. The added flexibility conferred by these structural units seems to be important for the binding to take place in the first instance, and for it to remain stable after it has occurred. It is possible that during extraction and isolation procedures, thiazolines in thiazoline containing lissoclinamides are oxidised to thiazoles. The formation of a CO_2 adduct of the dicopper complex of **10** again allows speculation of these dicopper complexes in the activation of dioxygen (see preceding paper).

5. General experimental

Compounds **9**, **10** were isolated and their spectra were assigned as described previously.[1]

5.1. Circular dichroism spectroscopy

CD spectra were acquired either at the BBSRC CD Facility at Stirling, UK or at the Faculty of Pharmacy, Universiteit Utrecht, Utrecht, the Netherlands. All spectra were recorded between 200 and 320 nm at a speed of 10 nm/min, resolution 0.2 nm, response 2 s and sensitivity 20 mdeg. All solutions were prepared in spectroscopic grade methanol. Metal solutions (0.1 M) were prepared from their chloride salts and then further diluted to give 0.5 and 1.0 mM stock solutions. Compounds **9**, **10** were prepared at 0.02 mg/mL in methanol.

5.2. Mass spectrometry

All mass spectrometry was carried out at Department of Biomolecular Mass Spectrometry, Utrecht University, Utrecht, the Netherlands. All samples were dissolved in HPLC grade methanol. Accurate mass measurements were carried out with a quadrupole-time-of-flight instrument (Q-TOF, Micromass) with nano-electrospray ionisation, at a resolution of 5000. An oligipeptide $(phenylalanine)_n$ was used as lock mass and *n* was chosen in such a way that the lock mass was higher than the accurate mass to be measured. All other MS experiments were carried out with an ion trap instrument (LC-Q Thermoquest/Finnigan) equipped with a nanoflow electrospray probe. The spray voltage was 0.9 kV and the capillary temperature 80°C. For MS^n experiments the relative collision energy was adjusted in such a way that the relative intensity of the parent ion was still approximately 30%.

5.3. Computational studies

Monte–Carlo conformational searching. The initial structures for compounds **9** and **10** were input by hand using the Macromodel v6.5 graphical user interface.[8] Chirality was assigned at all stereo centres but no further manual manipulation of the structures was undertaken, thus reducing operator bias in the conformational search as far as possible. The Macromodel native Monte–Carlo search algorithm was employed and all structures obtained were optimised using the MM2* force field.[9] To improve the comparability of the structures generated from this search with those obtained experimentally, the generalised Born solvent accessible area (GB/SA) continuum description of $CHCl_3$ was employed.

Investigation of hydrogen bonding. The resulting optimised structures obtained in the Monte–Carlo searches were refined using the semi-empirical PM3 molecular orbital method available in MOPAC v6.0[10] and the output geometry analysed for instances of short (<2.25 Å) inter-residue C═O…H–N distances.

Acknowledgements

L. A. M. is the recipient of an EPSRC quota award, and B. F. M. is the recipient of an EPSRC project grant. Financial support for the collection expedition came from the Carnegie Trust and the Nuffield Foundation. M. J. participated in a University of California, Santa Cruz marine natural products chemistry expedition partially supported by NIH grants CA47135 and CA52955. Hardware for the molecular modelling studies was purchased with a grant from the Royal Society. We wish to thank the BBSRC for support of the CD facility at Stirling. J. J. K. v. d. B., K. V. and A. J. R. H. thank the Dutch Science Foundation (NWO) for financial support.

References

1. Morris, L. A.; Kettenes-van den Bosch, J. J.; Versluis, K.;

Thompson, G. S.; Jaspars, M. *Tetrahedron* **2000**, *56*, 8345–8353.
2. Wipf, P.; Fritch, P. C.; Geib, S. J.; Sefler, A. M. *J. Am. Chem. Soc.* **1998**, *120*, 4105–4112.
3. Hawkins, C. J.; Lavin, M. F.; Marshall, K. A.; Brenk, A. L. v. d.; Watters, D. J. *J. Med. Chem.* **1990**, *33*, 1634–1638.
4. Rodger, A.; Norden, B. *Circular Dichroism and Linear Dichroism*, Oxford University: Oxford, 1997.
5. Schmitz, F. J.; Ksebati, M. B.; Chang, J. S.; Wang, J. L.; Hossain, M. B.; Helm, D. v. d.; Engel, M. H.; Serban, A.; Silfer, J. A. *J. Org. Chem.* **1989**, *54*, 3463–3472.
6. *Circular Dichroism and the Conformational Analysis of Biomolecules*, Fasman, G. D., Ed.; Plenum: New York, 1996.
7. Brenk, A. L. v. d.; Byriel, K. A.; Fairlie, D. P.; Gahan, L. R.; Hanson, G. R.; Hawkins, C. J.; Jones, A.; Kennard, C. H. L.; Moubaraki, B.; Murray, K. S. *Inorg. Chem.* **1994**, *33*, 3549–3557.
8. Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caulfield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.
9. Hasel, W.; Hendrickson, T. F.; Still, W. C. *Tetrahedron Comput. Meth.* **1988**, *1*, 103.
10. Dewar, M. J. S. *J. Mol. Struct.* **1983**, *100*, 41.

2 PERKIN

Conformational change in the thiazole and oxazoline containing cyclic octapeptides, the patellamides. Part 1. Cu^{2+} and Zn^{2+} induced conformational change †

Linda A. Morris,[a] Bruce F. Milne,[a] Gary S. Thompson[b] and Marcel Jaspars*[a]

[a] *Marine Natural Products Laboratory, Department of Chemistry, University of Aberdeen, Old Aberdeen, Scotland, UK AB24 3UE. E-mail: m.jaspars@abdn.ac.uk; Fax: +44 1224 272921; Tel: +44 1224 272895*

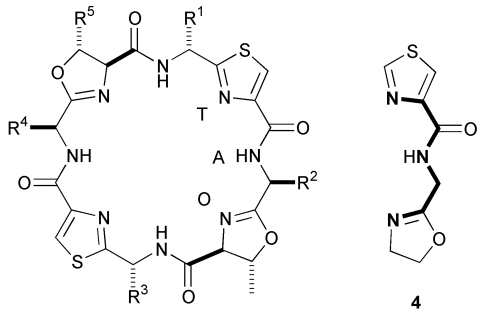
[b] *Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, UK LS2 9JT*

Received (in Cambridge, UK) 19th February 2002, Accepted 3rd April 2002
First published as an Advance Article on the web 24th April 2002

Conformational change during the binding of Cu^{2+} and Zn^{2+} to the thiazole and oxazoline containing cyclic octapeptides, the patellamides, is examined by a combination of experimental and theoretical methods. Circular dichroism and NOE-restrained molecular dynamics studies indicate that upon complexing with one equivalent of Cu^{2+}, patellamide C undergoes a change in conformation which pre-organises a second Cu^{2+} binding site, and that the binding of a second Cu^{2+} induces no further conformational change. The binding of Zn^{2+} induces little conformational change in patellamide C. A restrained conformational search shows that the conformational change induced by the addition of one equivalent of Cu^{2+} to patellamide C is an intrinsic design feature of the system. Electronic structure calculations indicate that the patellamides provide an ideal coordination environment for Cu^{2+}. On the basis of the evidence gathered, it can be proposed that Cu^{2+} is the biologically relevant metal for the patellamides.

Introduction

The thiazole (Thz) and oxazoline (Oxn) containing cyclic octapeptide patellamides (*e.g.* **1**–**3**), isolated from the Indo-Pacific ascidian (seasquirt) *Lissoclinum patella*, are known to bind both Cu^{2+} and Zn^{2+}.[1–5] The presence of thiazole and oxazoline rings has not yet been identified in other natural metal complexes or metalloproteins. One of these compounds, patellamide C (**1**), on complexing to one equivalent of Cu(II) forms a pre-organised binding site for a second Cu(II). Such a co-operative 'binary switching' ligand has not previously been described. The studies described below strongly implicate Cu(II) as the biologically relevant metal for this ligand.



4

1, patellamide C R^1=CH_2Ph; R^2=$CHMe_2$; R^3=Me; R^4=CHMeEt; R^5=Me
2, ascidiacyclamide, R^1=R^3=$CHMe_2$; R^2=R^4=CHMeEt; R^5=Me
3, patellamide A, R^1=R^3=$CHMe_2$; R^2=R^4=CHMeEt; R^5=H

The patellamide backbone can adopt either an open or closed conformation in solution depending on R^1–R^4 and solvent in the absence of metal ions.[6] If the side chains are symmetrically disposed (R^1 = R^3, R^2 = R^4; *e.g.* **2**, **3**) then it will adopt the 'open' conformation **B** in polar solvents (Fig. 1b) which has a positive maximum in the circular dichroism (CD) spectrum at 210 nm. An asymmetrical disposition of sidechains ($R^1 \neq R^3$, $R^2 \neq R^4$; *e.g.* **1**) results in the adoption of the 'closed' conformation **A** (Fig. 1a), which has a positive maximum at 250 nm in its CD spectrum.

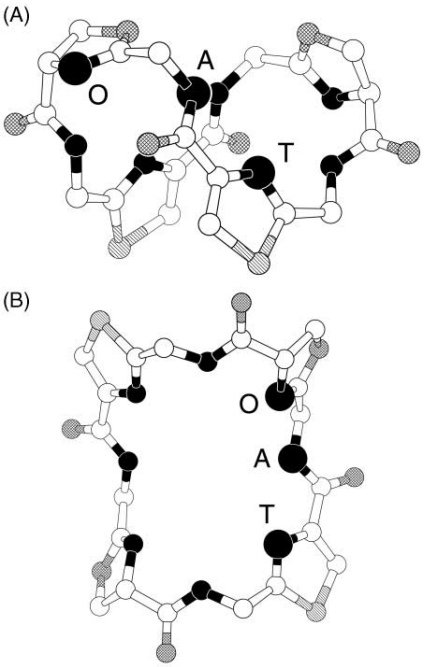


Fig. 1 Conformations of the patellamide skeleton indicating the 'TAO' motif. (**A**) The 'figure of eight' or folded conformation. (**B**) The 'square' or open conformation.

† Electronic supplementary information (ESI) available: further calculational details. See http://www.rsc.org/suppdata/p2/b2/b201823n/

 DOI: 10.1039/b201823n

Table 1 CD data for complexation of 0.025 mM **1** and **3** in MeOH with 0–2 equivalents of M^{2+} ($\Delta A/10^{-3}$ cm^{-1} mol^{-1})

	210 nm			250 nm		
	0 equiv.	1 equiv.	2 equiv.	0 equiv.	1 equiv.	2 equiv.
1–Cu^{2+}	0.185	0.226	0.230	0.151	0.014	0.012
1–Zn^{2+}	0.192	0.203	0.210	0.156	0.147	0.145
3–Cu^{2+}	0.183	0.224	0.223	0.017	0.009	0.006
3–Zn^{2+}	0.174	0.216	0.242	0.014	−0.016	−0.019

CD and mass spectrometric (MS) studies have shown that patellamide C (**1**) is selective for Cu(II) in the presence of Zn(II) and does not bind to Ni(II) or Co(II).[7] The present work deals with the difference in conformational behaviour between Cu(II) and Zn(II) binding of the patellamides. In the X-ray diffraction crystal structure of the dicopper carbonate complex of **2**,[3] the two Cu(II) ions bind to the nitrogens of the thiazole, amide and oxazoline groups which we have named the TAO motif (**4**). The Cu(II) adopts a distorted square pyramidal coordination, with the equatorial plane containing the TAO nitrogens and an O donor from the bridging carbonate and apical coordination from a water O. The TAO motif is pre-organised to accept Cu(II) in uncomplexed **2**, which adopts the open conformation (**B**) in solution and the crystalline state (Fig. 1).[8]

Results and discussion

Titrations of compound **3** with 0–2 equivalents of Cu(II) or Zn(II) and compound **1** with 0–2 equivalents of Zn(II) result in only very small changes in their CD spectra at both the 250 and 210 nm positive maxima indicating little change in conformation.[7] In combination with the MS data this shows the formation of a di-copper complex with the retention of the open conformation (**B**) for **3** (Table 1)[7] with 2 equivalents of Cu(II) and no complex formation and the retention of the closed conformation (**A**) in the case of **1** with Zn(II). This in stark contrast to the behaviour of **1** with Cu(II) where a switch from a closed conformation (**A**)[9] to an open conformation (**B**) occurs on binding to Cu(II); this has been confirmed by CD studies which show the presence of only two conformations.[7] These CD spectra indicate that addition of only 1 equivalent of Cu(II) to **1** (Table 1) causes the observed change in conformation. This conformational change pre-forms a second empty Cu(II) binding site, with the addition of a second equivalent of Cu(II) causing no further conformational change.

MS studies indicated a binding preference for the Ala(Thz)-Ile(Oxn) (left) TAO site in **1**.[7] To obtain the structure of **1** complexed with one equivalent of Cu(II), NOE restrained molecular dynamics calculations were used. A sample of **1** in CD_3OD was prepared in an inert atmosphere and one equivalent of Cu(II) was added as $CuCl_2$. Because of the fast relaxation caused by the Cu(II), selective 1D-NOE spectroscopy with short mixing times[10] was used (26 NOEs; only weak and medium correlations were observed due to the geometry of **1**). Most of the NOEs were found to be on the Val (right) side of **1**. This suggested that the Cu(II) was bound to the Ala(Thz)-Ile(Oxn) TAO (left) binding site with its paramagnetic effects making it impossible to detect the NOEs to that portion of the molecule. The 3D structure of **1**–Cu^{2+} was calculated using XPLOR 3.851;[11] during the calculations no specific Cu(II) binding model was used. The lowest energy ensemble consistent with the observed NOEs contained 18 structures, had only five minor NOE violations and an energy of 23.3 kcal mol^{-1}. The structures in the ensemble all had the same backbone conformation (RMSD = 0.00 ± 0.00 Å; Fig. 2, Table 2). When the structure of the complex **1**–Cu^{2+} was compared to the X-ray crystal structure of **2**,[8,12] an overlay of the TAO motif (**4**) on the Val (right) side was found to fit within typical experimental error (RMSD = 0.36 Å, Table 2, Fig. 3). This indicates that upon binding one equivalent of Cu(II), the structure of **1** changes from the closed conformation to the open conformation and in doing so makes a second binding site available for binding to another Cu(II).

Table 2 Distances in Å between thiazole (N_T), amide (N_A) and oxazoline (N_O) nitrogens in bound and free patellamides

Structure	Method	N_T–N_A	N_A–N_O	N_O–N_T
1–Cu^{2+}	NOE-MD[a]	2.84	2.93	5.04
1/restraints	MC[b]	2.90	2.73	4.71
2	XRD[c]	2.82	2.79	4.80
2–Cu_2CO_3	XRD[c]	2.59	2.64	4.01
4–$Cu^{2+}(MeOH)_2$	*ab initio*[d]	2.66	2.62	4.14
4–$Cu^{2+}(H_2O)_2$	*ab initio*[d]	2.65	2.62	4.11

[a] NOE restrained molecular dynamics calculations. [b] Monte-Carlo conformational search. [c] Published X-ray crystal structure. [d] *ab initio* Electronic structure calculations.

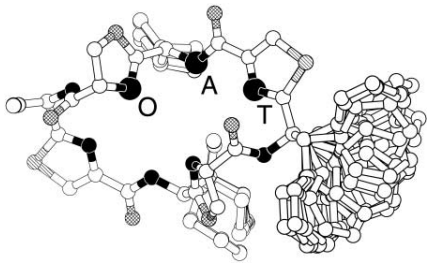


Fig. 2 A cluster of 18 low energy structures of **1**–Cu^{2+} indicating the TAO motif.

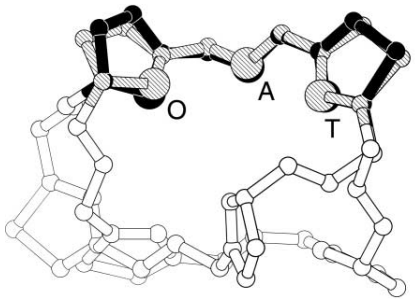


Fig. 3 Overlay of the minimum energy conformation of **1**–Cu^{2+} (black) and the crystal structure of ascidiacyclamide **2** (hatched).

NOE restrained molecular dynamics calculations on **1** with 2 equivalents of Zn(II) in CD_3OD, using 25 NOE restraints gave a structure with one NOE violation, an energy of 28.4 kcal mol^{-1} and a backbone RMSD of 0.00 ± 0.00 Å for the 38 lowest energy structures (Fig. 4). In qualitative terms, after addition of 2 equivalents of Zn(II), **1** appears to adopt a slightly expanded version of the closed conformation **A** (Fig. 5); however, these changes are maybe within experimental error. Previous studies have suggested that Zn(II) may complex with the patellamides, but the present study suggests that little conformational change is involved.[4]

A Monte-Carlo conformational search of the coordinate space for **1** in which the Ala(Thz)-Ile(Oxn) TAO site was restrained into the copper-bound conformation was conducted

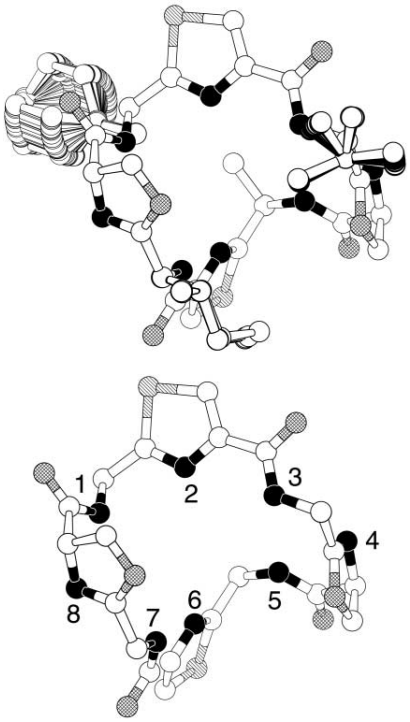


Fig. 4 Ensemble of the 38 lowest energy structures of Zn(II) bound **1** (top) and the minimum energy structure showing residue numbers for clarity (bottom).

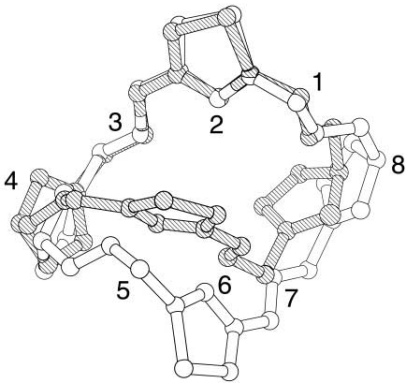


Fig. 5 Overlay of Zn(II) bound and uncomplexed (hatched) **1** (heavy atoms only) with residue numbers shown at the amide nitrogen for clarity. The models were overlaid using the backbone atoms of the three residues which gave the lowest RMSD (1, 2 and 3; RMSD = 0.20 Å).

using MacroModel.[13] Again, no specific Cu(II) binding model was employed, so that the effect on Phe(Thz)-Val(Oxn) of restraining Ala(Thz)-Ile(Oxn) in the Cu(II) bound form could be studied. The lowest energy conformation found during the search (5000 steps) again showed pre-organisation of the Phe-(Thz)-Val(Thz) TAO site. This showed good overlay of the TAO motif (**4**) on the Val (right) side of the structure with the crystal structure determined for **2** (RMSD = 0.09 Å, Table 2). The conformational search indicates that the pre-organisation of a second TAO binding site is an integral conformational feature of the patellamides.

ab initio Electronic structure calculations were performed at the ROHF level on **4**–Cu(II) using the PC GAMESS version[14] of the GAMESS (US) QC package.[15] The SBKJC effective core potential basis set with additional heavy atom polarisation functions was employed for all calculations.[16–18] The initial structures had the metal in an octahedral coordination geometry composed of the three co-planar TAO nitrogens with the remaining sites occupied by the oxygen atoms of 3 solvent molecules (H_2O and MeOH to simulate both the natural and

Table 3 Selected distances and angles for structures from electronic structure calculations in comparison to those determined from a published X-ray crystal structure

	2–Cu_2CO_3[a]	**4**–$Cu^{2+}(MeOH)_2$	**4**–$Cu^{2+}(H_2O)_2$
Angles/deg			
N_T–Cu–N_A	80.7	81.7	82.2
N_O–Cu–N_A	81.3	81.6	82.1
Distances/Å			
N_T–N_A	2.590	2.659	2.652
N_A–N_O	2.640	2.619	2.616
N_O–N_T	4.069	4.141	4.109
N_T–Cu	2.101	2.119	2.098
N_A–Cu	1.925	1.940	1.931
N_O–Cu	2.052	2.066	2.051

[a] Published X-ray crystal structure.[3]

experimental conditions experienced by the metal). During the course of both optimisations the solvent ligand below the TAO plane was ejected resulting in a distorted square pyramidal coordination environment (Fig. 6). The results from the

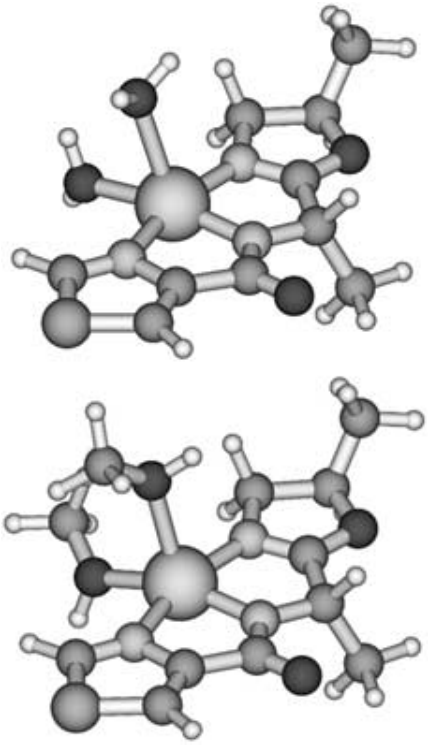


Fig. 6 Output geometry from electronic structure calculations for **4**–$Cu^{2+}(H_2O)_2$ (top) and **4**–$Cu^{2+}(MeOH)_2$ (bottom).

calculation in MeOH showed good overlap with the TAO motif in the X-ray crystal structure of **2**–Cu_2CO_3 (RMSD = 0.085 Å, Table 2), indicating that Thz-(C=O)–NH–CH_2-Oxn provides the ideal geometry for complexing to Cu(II). The Cu(II) geometry calculated is almost identical to that obtained from the crystal structure.[3] Selected bond angles and distances for **4** complexed with Cu^{2+} in H_2O and MeOH are given in Table 3 in comparison to those measured from the crystal structure for **2**–Cu_2CO_3[3] showing excellent correspondence between experimental and theoretical values.

Conclusions

The above studies suggest that the conformational change observed on the addition of one equivalent of Cu^{2+} to patellamide C, creating a second Cu^{2+} binding site, is an intrinsic

design feature of these compounds. Electronic structure calculations show that the coordination environment present in these molecules is ideal for Cu^{2+}. It can therefore be proposed that Cu(II) is the biologically relevant metal in this unusual natural product. In addition, the binary switching nature of the patellamides on complexing to Cu(II) could prove to be useful in the creation of novel molecular devices.

Experimental

NOE restrained molecular dynamics

Restraints were derived from T-ROESY[19] or 1D selective NOE[10] spectra and classified as weak, medium or strong. ROE's were quantified by contour counting. Restrained molecular dynamics calculations were carried out with XPLOR 3.851[11] with a force field with repulsive non-bonded terms. No dielectric term was included in the calculations to take account of solvent effects. *ab initio* Simulated annealing calculations (YASAP 3.0: 120 ps total time simulated annealing from 2000 to 100 K, 200 steps minimisation)[20,21] were used to calculate structures from a starting conformation with randomised ϕ and ψ angles. Sum averaging[22] was used for all methyl, methylene, and aromatic ring atom pairs. Assignment of prochiral groups was achieved by floating assignment and swapping of prochiral groups.[23,24] A reduced set of non-bonded interactions and a reduced representation of the amino acid side chains was used during the conformation search phase.[25] During all stages of the simulated annealing calculation the temperature was maintained by coupling to a heat bath[26] with a coupling frequency of 10 ps^{-1}. From each ensemble the lowest energy structures were refined using a simulated annealing with slow cooling protocol (600 ps, cooling from 1500 to 100 K, 4000 steps of minimisation). The lowest energy structures from each ensemble were selected to represent their structures. The overlay and display of structures was achieved using Molmol.[27] Figures were created using DINO[28] and rendered using POV-Ray™.[29]

Monte-Carlo conformational searching

The input structures were drawn in by hand and the peptide backbone kept planar until the beginning of the run. Restraints were added (as described in the electronic supplementary information) to simulate the presence of a bound Cu(II). 5000 steps of Monte-Carlo conformational searching were followed by (up to) 5000 energy minimisation steps within the MM2* force field.[13] The generalised Born solvent accessible area (GB/SA) continuum solvent model was used to simulate the chloroform environment in which the NMR values were obtained.[30] Cut-offs for non-bonded interactions were employed (12 Å for electrostatics and 7 Å for van der Waals interactions) giving a distance-dependent dielectric environment for the solute molecule.

Acknowledgements

BFM is the recipient of an EPSRC PhD award. LAM and GST were recipients of BBSRC PDRA awards (1/E12737 and 49/SBD07527) The Carnegie Trust provided funds for a 6 week visit to Steve Homans' laboratory in Leeds to carry out MD studies.

References

1 C. J. Hawkins, *Pure Appl. Chem.*, 1988, **60**, 1267–1270.
2 A. L. van den Brenk, D. P. Fairlie, G. R. Hanson, L. R. Gahan, C. J. Hawkins and A. Jones, *Inorg. Chem.*, 1994, **33**, 2280–2289.
3 A. L. van den Brenk, K. A. Byriel, D. P. Fairlie, L. R. Gahan, G. R. Hanson, C. J. Hawkins, A. Jones, C. H. L. Kennard, B. Moubaraki and K. S. Murray, *Inorg. Chem.*, 1994, **33**, 3549–3557.
4 L. Grondahl, N. Sokolenko, G. Abbenante, D. P. Fairlie, G. R. Hanson and L. R. Gahan, *J. Chem. Soc., Dalton Trans.*, 1999, 1227–1234.
5 D. J. Freeman, G. Pattenden, A. F. Drake and G. Siligardi, *J. Chem. Soc., Perkin Trans. 2*, 1998, 129–135.
6 M. Doi, F. Shinozaki, Y. In, T. Ishida, D. Yamamoto, M. Kamigauchi, M. Sugiura, Y. Hamada, K. Khoda and T. Shioiri, *Biopolymers*, 1999, **49**, 459–469.
7 L. A. Morris, M. Jaspars, J. J. Kettenes van den Bosch, K. Versluis, A. J. R. Heck, S. M. Kelly and N. C. Price, *Tetrahedron*, 2001, **57**, 3185–3197.
8 T. Ishida, M. Tanaka, M. Nabae, M. Inoue, S. Kato, T. Hamada and T. Shioiri, *J. Org. Chem.*, 1988, **53**, 107–112.
9 T. Ishida, Y. In, F. Shinozaki, M. Doi, D. Yamamoto, Y. Hamada, T. Shioiri, M. Kamigauchi and M. Suguira, *J. Org. Chem.*, 1995, **60**, 3944–3952.
10 D. Uhrín and P. N. Barlow, *J. Magn. Reson.*, 1997, **126**, 248–255.
11 A. T. Brünger, X-PLOR: A system for X-ray crystallography and NMR, Version 3.1, Yale University, New Haven, 1993.
12 Y. Hamamoto, M. Endo, M. Nakagawa, T. Nakanishi and K. Mizukawa, *J. Chem. Soc., Chem. Commun.*, 1983, 323–324.
13 F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caulfield, G. Chang, T. Hendrickson and W. C. Still, *J. Comput. Chem.*, 1990, **11**, 440–467.
14 A. A. Granovsky, http://classic.chem.msu.su/gran/gamess/index.html
15 M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis and J. A. Montgomery, *J. Comput. Chem.*, 1993, **14**, 1347–1363.
16 J. S. Binkley, J. A. Pople and W. J. Hehre, *J. Am. Chem. Soc.*, 1980, **102**, 939–946.
17 W. J. Stevens, H. Basch and M. Krauss, *J. Chem. Phys.*, 1984, **81**, 6026–6033.
18 W. J. Stevens, M. Krauss, H. Basch and P. G. Jasien, *Can. J. Chem.*, 1992, **70**, 612–630.
19 A. Bax and D. G. Davis, *J. Mag. Reson.*, 1985, **63**, 207–213.
20 M. Nilges, A. M. Gronenborn, A. T. Brünger and G. M. Clore, *Protein Eng.*, 1988, **2**, 27–38.
21 M. Nilges, J. Kuszewski and A. T. Brünger, in *Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance*, J. C. Hoch, F. M. Poulsen and C. Redfield (Eds.), Plenum Press, New York, 1991, Vol. 225, pp. 451–455.
22 A. T. Brünger, G. M. Clore, A. M. Gronenborn and M. Karplus, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, **83**, 3801–3805.
23 T. A. Holak, M. Nilges and H. Oschkinat, *FEBS Lett.*, 1989, **242**, 218–224.
24 R. H. A. Folmer, C. W. Hilbers, R. N. H. Konings and M. Nilges, *J. Biomol. NMR*, 1997, **9**, 245–258.
25 M. Nilges, *Proteins: Struct., Funct., Genet.*, 1993, **17**, 297–309.
26 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.
27 R. M. B. Koradi and K. Wuthrich, *J. Mol. Graphics*, 1994, **14**, 51–59.
28 DINO: Visualising Structural Biology, A. Philippsen, Basel, Switzerland, 2001; http://www.bioz.unibas.ch/~xray/dino.
29 POV-Ray Team™, Persistence of Vision™ Ray Tracer (POV-Ray™), Version 3.1g, Williamstown, Australia, 1999.
30 W. Hasel, T. F. Hendrickson and W. C. Still, *Tetrahedron Comput. Methodol.*, 1988, **1**, 103–116.

Conformational change in the thiazole and oxazoline containing cyclic octapeptides, the patellamides. Part 2. Solvent dependent conformational change †

Bruce F. Milne,[a] Linda A. Morris,[a] Marcel Jaspars *[a] and Gary S. Thompson [b]

[a] *Marine Natural Products Laboratory, Department of Chemistry, University of Aberdeen, Old Aberdeen, Scotland, UK AB24 3UE. E-mail: m.jaspars@abdn.ac.uk; Fax: +44 1224 272921; Tel: +44 1224 272895*

[b] *Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, UK LS2 9JT*

Received (in Cambridge, UK) 19th February 2002, Accepted 3rd April 2002
First published as an Advance Article on the web 24th April 2002

Solvent dependent conformational change of the thiazole and oxazoline containing cyclic peptides, the patellamides, is examined by a combination of experimental and theoretical methods. A mechanism for the simultaneous formation of two type-II β-turns in the patellamides is proposed based on molecular dynamics and NOE restrained molecular dynamics studies as well as literature evidence. The effect of the solvent and desymmetrisation of the patellamides is crucial, with symmetrical patellamides in polar solvents giving the open type-I conformation, whereas symmetrical patellamides in non-polar solvents and asymmetrical patellamides in both polar and non-polar solvents give rise to the folded type-II conformation.

Introduction

The unusual thiazole and oxazoline containing cyclic octapeptide patellamides (**1**–**6**, Scheme 1, Table 1), isolated from the ascidian (seasquirt) *Lissoclinum patella*, have been much investigated due to their interesting biological activity, such as potent cytotoxicity and the reduction of multi-drug resistance of certain types of lymphoblasts.[1] They can be divided into two naturally occurring families: symmetrical patellamides where $R^1 = R^3$ and $R^2 = R^4$ (**1**, **2**) and asymmetrical patellamides where $R^1 \neq R^3$ and R^2 may equal R^4 (**3**–**6**) and $R^1 = CH_2Ph$. The patellamides exist in one of two conformations depending on the side-chains present in the molecule, a 'square' conformation (I) and a 'figure of eight' conformation (II) (Scheme 1).[2]

† Electronic supplementary information (ESI) available: further calculational details. See http://www.rsc.org/suppdata/p2/b2/b201824c/

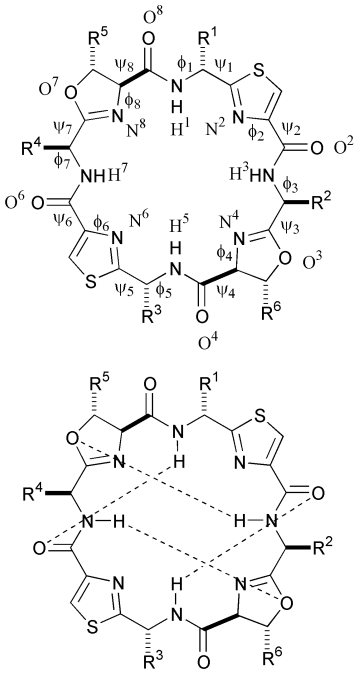
Direct NOE and CD evidence for the solution state conformations of various natural patellamides and a number of synthetic analogues are known (Table 1), and from the available data a clear trend emerges. The symmetric compounds such as patellamide A (**2**) and the synthetic patellamide analogue **7** with $R^1 = R^3$; $R^2 = R^4$ all adopt a square conformation (I) in methanolic solution at room temperature. This can be contrasted with patellamides B (**3**), C (**4**) and E (**6**)[2–4] and the synthetic analogues **8**–**10** ($R^1 \neq R^3$; R^2 may equal R^4) which all take up the closed 'figure of eight' conformation (II) in both polar and non-polar solvents. The effect of the side-chains in these compounds on their conformation is corroborated by the X-ray diffraction crystal structures of the symmetrical ascidiacyclamide (**1**) and its synthetic analogue **7** which both adopt the open structure (I). The X-ray crystal structure of the asymmetric variant, patellamide D (**5**), is in the closed conformation (II).

Table 1 Structures and published solution and crystalline state conformations of natural and synthetic patellamides **1**–**10**

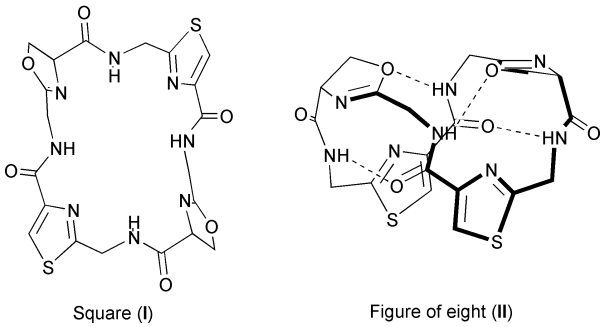
Compound	R^1	R^2	R^3	R^4	R^5	R^6	Conformation	Method	Solvent	Ref.
1 Ascidiacyclamide	*i*-Pr	*s*-Bu	*i*-Pr	*s*-Bu	Me	Me	I	Crystal	C_6H_6	6
							I	NMR	$CDCl_3$	6
2 Patellamide A	*i*-Pr	*s*-Bu	*i*-Pr	*s*-Bu	H	Me	I	Crystal	MeOH	12
							I	CD (299 K)	MeOH	4
							II	CD (199 K)	MeOH	4
							I	CD	MeOH	This study
							II	NMR-MD	$CDCl_3$	This study
3 Patellamide B	CH_2Ph	CH_2CHMe_2	Me	*s*-Bu	Me	Me	II	NMR-MD	$CDCl_3$	2
							II	CD	MeOH	4
4 Patellamide C	CH_2Ph	*s*-Bu	Me	*i*-Pr	Me	Me	II	NMR-MD	$CDCl_3$	2
							II	CD	MeOH	This study
5 Patellamide D	CH_2Ph	*i*-Pr	Me	*i*-Pr	Me	Me	II	Crystal, MD	MeOH	13
6 Patellamide E	CH_2Ph	*i*-Pr	*i*-Pr	*s*-Bu	Me	Me	II	CD	MeOH	4
7[a]	*s*-Bu	*i*-Pr	*s*-Bu	*i*-Pr	H	H	I	CD, NMR	MeOH	14
8[a]	*i*-Pr	H	*i*-Pr	*s*-Bu	Me	Me	II	NMR-MD	DMSO	3
9[a]	*i*-Pr	CH_2CHMe_2	*i*-Pr	*s*-Bu	Me	Me	I	Crystal	C_6H_6	3
							I/II	NMR-MD	DMSO-d_6	3
10[a]	*i*-Pr	CH_2Ph	*i*-Pr	*s*-Bu	Me	Me	II	NMR-MD	DMSO-d_6	3
11[b]	CH_2Ph	*s*-Bu	*i*-Pr	*s*-Bu	Me	Me				

[a] Denotes synthetic analogue. [b] Used in theoretical calculations only.

 DOI: 10.1039/b201824c



H-bonds monitored during molecular dynamics.



Square (I) Figure of eight (II)

Scheme 1 The chemical structure of the patellamides, the two conformations and the hydrogen bonds monitored during molecular dynamics simulations.

Studies by Doi *et al.* show that disturbing the symmetry at R^2 also alters the conformation.[3] Changing R^2 from *s*-Bu (**1**) in conformation (I) to H (**8**) results in the closed conformation (II) whereas altering it to CH_2CHMe_2 (**9**) gives the open conformation (I) in the crystalline state, but results in a hybrid I/II conformation in DMSO-d_6 solution. Changing the *s*-Bu at R^2 for CH_2Ph (**10**) results in the closed conformation (II) in solution.

Variable temperature circular dichroism studies[4] and our own molecular mechanics calculations suggest that the closed conformation (II) is favoured by enthalpic (ΔH) contributions to the free energy of the molecule. This raises the important question as to why the 'square' or 'open' conformation (I) exists almost exclusively at room temperature for symmetrical **1** (R^1 = *i*-Pr) and the pseudosymmetrical patellamide A (**2**, R^1 = *i*-Pr), whereas the 'figure of eight' or 'closed' (II) conformation is dominant for **3**–**6** (R^1 = CH_2Ph). Evidence is presented below for the importance of solvent polarity on the conformational state and a mechanism is proposed for the interconversion of conformation I to II.

Results and discussion

It is obvious from the three dimensional structure of the open (I) and closed (II) form of the patellamide structure that large changes in ϕ and ψ, the backbone torsion angles, are involved in their interchange (Table 2), with the largest changes being

Table 2 ϕ and ψ torsional angle values (deg) and changes in ϕ and ψ for conformations I and II

	1 (square, I)		**5** (figure of eight, II)		I ⟶ II	
Residue	ϕ	ψ	ϕ	ψ	$\Delta\phi$	$\Delta\psi$
1	134.5	47.0	83.8	−12.4	50.7	59.4
2	179.0	−8.0	179.1	−4.0	−0.1	−4.0
3	−127.0	−21.2	−76.4	121.1	−50.6	−142.3
4	113.7	18.8	101.5	−14.3	12.2	33.1
5	134.5	47.0	91.0	−11.1	43.5	58.1
6	179.0	−8.0	179.5	−4.9	−0.5	−3.1
7	−127.0	−21.2	−80.2	113.6	−46.8	−134.8
8	113.7	18.8	102.0	−1.4	11.7	20.2

Table 3 Type-II β turn in the figure of eight conformation (II) for compound **5**[5]

	ϕ_i/deg	ψ_i/deg	ϕ_{i+1}/deg	ψ_{i+1}/deg
Ideal	−60.0	120.0	80.0	0.0
Residue 3–4	−76.4	121.1	101.5	−14.3
Residue 7–8	−80.2	113.6	102.0	−1.4

observed in ψ_3 and ψ_7. The process is driven by the formation of two unusual type-II β turns[5] (residues 3–4 and 7–8) in the closed conformation (II). These turns are stabilised in (Table 3) by the formation of two new C=O to N–H hydrogen bonds (O^2–H^5 and O^6–H^1) and two new oxazoline O to N–H hydrogen bonds (O^3–H^7 and O^7–H^3), which are not part of normal peptidic β-turns (Scheme 1). This conformation is not static and dynamics simulations show that at any given moment the four hydrogen bonds are not present simultaneously; one may be formed as another is broken. This is consistent with studies by Ishida *et al.* which show that all N–H protons are exposed to solvent for **3** and **4** in the conformation II.[3]

Little information is available relating to the conformations of the symmetrical patellamides in non-polar solvents except for the early work by Ishida *et al.*[6] in which H_N–N–C_α–H_α NMR coupling constants measured for ascidiacyclamide (**1**) in $CDCl_3$ solution gave torsional angles consistent with the crystal structure of the type-I conformation. However, our measurement of the same coupling constants for patellamide C (**4**), which exists in the closed conformation (II) in $CDCl_3$ solution,[2] gave similar values to those obtained for **1**, suggesting that these coupling constants are not a good indicator of conformation. This ambiguity may result from a combination of two factors: a rapid averaging of the torsional angles, and the fact the coupling constants are on a steep part of the Karplus curve. Furthermore, it is not possible to distinguish between the open (I) and closed (II) form of the patellamides using CD spectra in $CHCl_3$ due to solvent interference. It was therefore decided to investigate the conformations of the pseudosymmetrical patellamide A (**2**) and the asymmetrical patellamide C (**4**) in pure $CDCl_3$ by NOE restrained molecular dynamics with the conformations present in MeOH being determined by circular dichroism.

The conformations of the open (I) and closed (II) forms of the patellamides each give a characteristic circular dichroism (CD) spectrum with positive maxima at 211 and 205 nm for the first form and 250 nm for the second.[7] Therefore, the presence of a positive maximum at 250 nm for patellamide C can be used to indicate the presence of the closed conformation (II) and its absence for patellamide A suggests the presence of the open conformation (I).

NOE restrained molecular dynamics has been used previously to determine the solution conformation of **3**, and **4**[2] and the synthetic analogues **8**–**10**.[3] During the course of this study the structure of **4** in $CDCl_3$ was obtained from approximately 60 NOEs and our calculation of the conformation is in

good agreement with that calculated previously (backbone RMSD 0.20 ± 0.09 Å).[2]

Patellamide A (**2**) is nearly symmetrical and only approximately 30 NOEs could be obtained due to chemical shift degeneracy. This degeneracy can prove to be a major problem in restrained molecular dynamics calculations as if all possible restraints are applied many violations will occur and this will lead to strained and distorted structures. Therefore, to solve this problem we appropriated the technique of ambiguous distance restraints used with biopolymers for the calculation of dimer structures and in automated structure calculation protocols. Thus, for each possible assignment a function based on the $1/r^6$ sum average of the two possible distances was used as the restraining function during X-PLOR molecular dynamics calculation.[8] As an example, the strong restraint from $3H_\alpha$ to NH^5 was set as $3H_\alpha$ to NH^5 *or* $3H_\alpha$ to NH^1 as the shifts of NH^5 and NH^1 are isochronous in the ^1H NMR spectrum. This ambiguous restraint allows both assignments to contribute in an $1/r^6$ sum average, but at the end of the calculation it is clear that this restraint is satisfied by a short distance between $3H_\alpha$ to NH^5 with the distance $3H_\alpha$ to NH^1 contributing little to the restraining energy. A total of 52 low energy structures were calculated from 100 structures in which ϕ and ψ had been randomised with 8 *C*=O to 1 *N*–H being set as an 'elastic bond' to relieve strain in the ring. All of these structures were refined and a cluster of 22 structures with the same backbone geometry and a low target energy function were chosen as representative of the solution structure (Fig. 1, backbone RMSD 0.22 ± 0.14 Å,

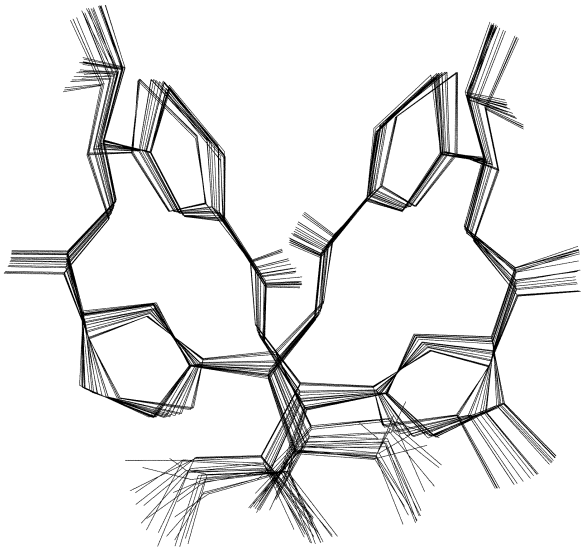


Fig. 1 Ensemble of the 22 minimum energy structures of patellamide A (**2**) in $CDCl_3$ (backbones only shown).

E_{Total} = 11.8 kcal mol^{-1}). The final conformation is closed (II), and the agreement with our calculated structure for **4** is good (Fig. 2, backbone RMSD = 0.34 Å).

These results indicate that both symmetrical and asymmetrical patellamides can adopt the closed conformation (II) in non-polar solvents. To test the effect of solvent polarity on the structure of the patellamides we carried out molecular mechanics minimisations[9] of ascidiacyclamide (**1**), and a non-natural ascidiacyclamide variant **11** in which $R^1 = CH_2Ph$. These were used to represent the symmetrical and asymmetrical patellamides respectively, changing R^1 from *i*-Pr to CH_2Ph whilst the rest of the molecule remained constant. Energies were calculated for each structure in both the open (I) and closed states (II) using a GB/SA continuum treatment of solvation in MacroModel.[10] The solvent parameters for H_2O and $CHCl_3$ were used. This system provided a pair of solvents that emphasise the effect of changing from a non-polar aprotic solvent to a polar protic solvent and uses a force field parametrisation which is readily available.[10] The results show that the folding from the open (I) to closed (II) form was highly favoured enthalpically in all cases. In H_2O $\Delta E_{I \longrightarrow II}$ was −12 kJ mol^{-1} for **1**, −10 kJ mol^{-1} for **11**, whereas in $CHCl_3$ the figures were −58 and −66 kJ mol^{-1} respectively. This indicates that, within the bounds of error, purely based on enthalpic terms including electrostatic and van der Waals effects the closed form is always favoured. This is consistent with the observations of Freeman *et al.*[4] who observed by CD that at 'low temperature' (173 K) in methanol patellamide A has a closed conformation (II) but that at 'high temperature' (298 K) it is open (II).

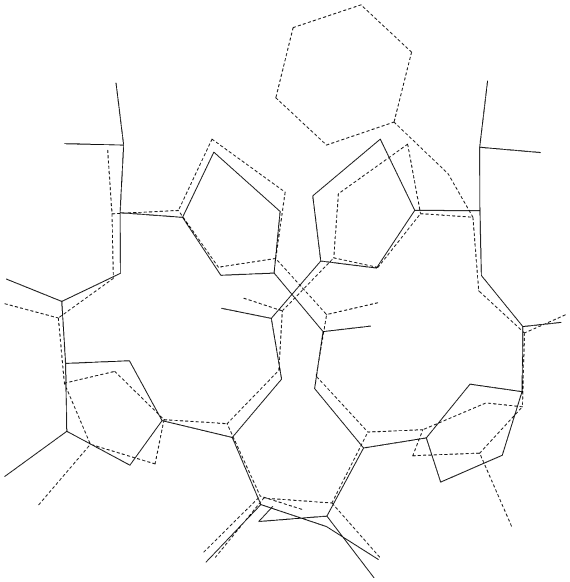


Fig. 2 Overlay of the minimum energy structure of patellamide A (**2**, line) and the minimum energy structure of patellamide C (**4**, dotted line) in $CDCl_3$ (backbones only shown).

To investigate the process by which the open and closed forms interconvert, molecular dynamics (MD) simulations were carried out on **1** and **11** at 300 K in H_2O and $CHCl_3$. The same GB/SA treatment of solvation as described above for the determination of enthalpy was used. The starting point for each MD simulation was a type-I conformation selected at random from the output of a previous 10 ps run. Six MD simulations were conducted for each system with a 1 fs time-steps for 200 ps and 1000 structures were sampled. For these structures the O^2–H^5, O^6–H^1, O^3–H^7 and O^7–H^3 distances were used to monitor the conversion of the structure from the open (I) to closed form (II) (Fig. 3). It was assumed that if these distances fell below 2.5 Å a hydrogen bond was able to form. Fig. 3 gives a good qualitative description of the evolution of the interconversion process with time. The results for **1** and **11** in $CHCl_3$ agree with the experimental observation that both symmetric and asymmetric patellamides will fold to give the closed conformation (II) in non-polar solvents. In polar solvents ascidiacyclamide (**1**) is expected to remain in the open conformation (I) (*cf.* **2** in MeOH, Table 1) which is confirmed by the aqueous theoretical simulation. The results of the MD simulations for **11** in H_2O however are more ambiguous as both open and closed forms are observed, whereas CD studies of **4** in MeOH show it is in the closed form (II). However, a number of factors must be considered, as the distance–time graph for **11** in H_2O (Fig. 3) shows that the monitored distances show significant excursions to near hydrogen bonding lengths during many of the simulations. Furthermore, the solvent system in the simulation used has a much higher dielectric (78.5) than the polar solvents used in the experimental studies which resulted in the closed (II) conformation (MeOH, 32.7; DMSO, 46.7). It could be postulated that this may tend to over emphasise the open conformation (I) in the simulations.

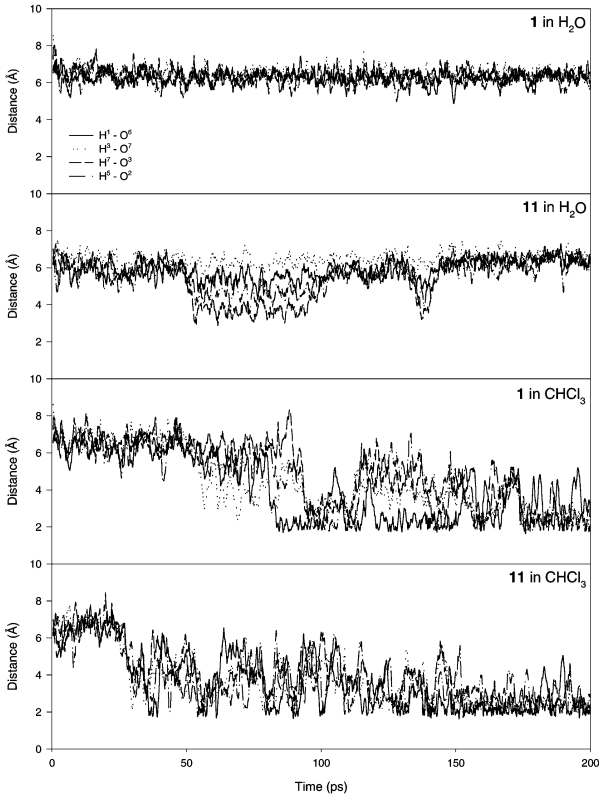


Fig. 3 Monitoring four key H-bonding distance during molecular dynamics simulations of ascidiacyclamide (**1**) and **11** (R^1 = CH_2Ph, R^2 = *s*-Bu, R^3 = *i*-Pr, R^4 = *s*-Bu) in H_2O and $CHCl_3$ at 300 K.

The MD simulations for **1** and **11** in $CHCl_3$ were used to construct a qualitative picture of how the folding from the open (I) to closed conformations (II) occurs; the observed pathway is different from the empirical model proposed by Ishida *et al.*[2] Fig. 4 shows the process broken up into four distinct steps, although in the MD simulations adjacent steps may occur in a concerted manner. The first step in almost all simulations in

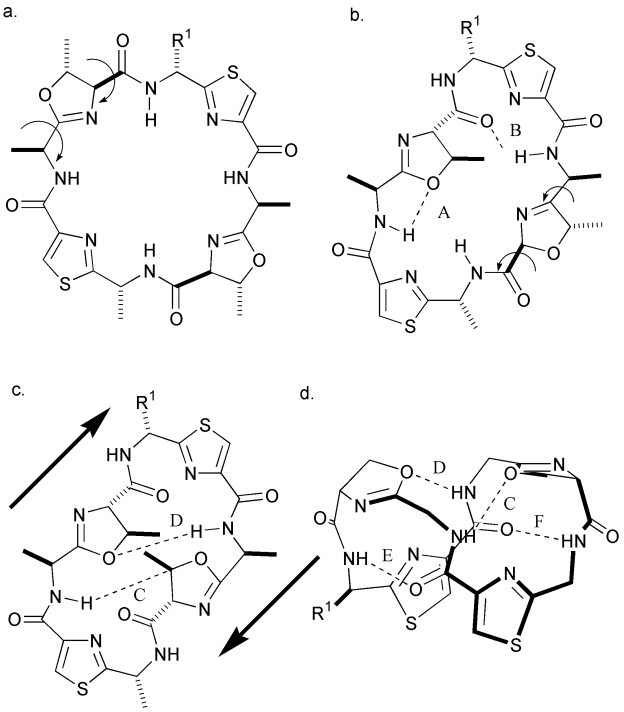


Fig. 4 A proposed mechanism for the folding process I ⟶ II derived from molecular dynamics simulations.

which folding occurred was the rotation about one oxazoline ring as in Fig. 4a. This event was monitored using the distances O^3–H^3 and O^7–H^7 which are roughly 4 Å apart in the open conformation (I) and only 2.5–3.0 Å upon rotation of the oxazoline ring (Fig. 5). Initially it was thought that this rotation

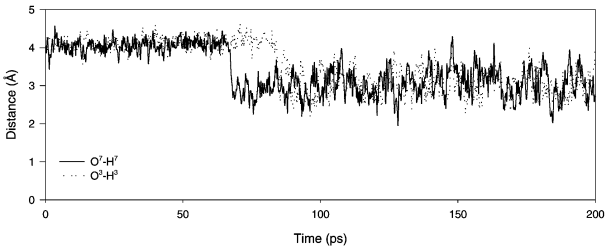


Fig. 5 Monitoring O^3–H^3 and O^7–H^7 distances during a molecular dynamics simulation of **11** in $CHCl_3$ at 300 K.

was facilitated by the presence of a symmetry disturbing group at R^1, but in both **1** and **11** in $CHCl_3$ the first rotation could occur at either oxazoline ring. This rotation may facilitate the formation of a hydrogen bond from the oxazoline oxygen to the adjacent NH (H-bond A in Fig. 4b) and a stabilising hydrogen bond B (Fig. 4b). The next step is the rotation of the second oxazoline ring, which is in some cases simultaneous with the rotation of the first oxazoline ring, giving rise to the structure in Fig. 4c in which distances C and D are of the order of 3.7 Å. In order to go to the closed conformation (II) shown in Fig. 4d, the two halves of the molecule need to experience 'shear' as indicated by the arrows to bring C and D into hydrogen bonding distance. This is corroborated by the observation that lengths of C and D are inversely related to the length of O^8–H^3 (H-bond B in Fig. 4b) and O^4–H^7. The 'shear' process simultaneously brings about the formation of hydrogen bonds E and F, thus completing the formation of two type-II β-turns.

Conclusions

It is clear from these studies that asymmetry in the structure of patellamides favours conformational change to the 'closed' conformation II. The molecular dynamics simulations predict the experimentally observed conformation correctly except in the case of the asymmetrical variant **11** in polar solvents. This may be due to the large increase in dielectric in going from MeOH, in which the CD studies are conducted, to H_2O, in which the molecular dynamics simulations are performed. This may be solved in future studies by developing a continuum model of MeOH. Separate simulations of **11** in H_2O starting from conformation II do not show the expected change back to conformation I, suggesting that the use of explicit solvent molecules might be desirable in future simulations.

This study gives an indication of the design features of this intriguing natural product allowing it to achieve such a fold. Certain motifs from this molecule may be used to study protein folding. Planned studies include this, as well as investigation of the above process by ultrafast spectroscopic measurements using temperature jumps and stopped-flow methods.

Experimental

Calculation of structures

30 NOE restraints were derived from the T-ROESY spectrum (T_{mix} = 400 ms) and classified as weak (1.8–5.0 Å), medium (1.8–3.5 Å) or strong (1.8–2.5 Å). ROE's were quantified by contour counting. Dihedral restraints were included for the four H_N–N–C_α angles. Restrained molecular dynamics calculations were carried out with X-PLOR 3.851[8] with a force field with repulsive non-bonded terms. No dielectric term was included in the calculations to take account of solvent effects.

ab initio Simulated annealing calculations (YASAP 3.0: 120 ps total time simulated annealing from 2000 to 100 K, 200 steps minimisation) were used to calculate structures from 100 starting conformations with randomised ϕ and ψ torsion angles using an elastic bond between 8 *C*=O and 1 *N*–H to relieve strain. From this ensemble 52 structures were refined using a simulated annealing with slow cooling protocol (600 ps, cooling from 1500 to 100 K, 4000 steps of minimisation). The 22 lowest energy structures from the ensemble were selected to represent the final structure. The overlay and display of structures was achieved using Molmol.[11]

Molecular dynamics

Molecular dynamics simulations were performed at 300 K over a 200 ps time period using the MM2* force field within Macromodel v6.5.[9] Starting structures were selected at random from the output of previous 10 ps equilibration runs. The simulations were performed in solvent (H_2O and $CHCl_3$) described by the GB/SA continuum model.[10] The update of the non-bonded interactions occurring within the system was performed with a period of 1 fs and cut-off distances were fixed at 12 and 7 Å for electrostatic and van der Waal's interactions, respectively. These parameters allowed for a higher degree of accuracy in the monitoring of any event that involved rapid and/or large conformational changes. The simulations in each solvent were repeated 6 times using different starting structures in an attempt to reduce the possibility that the initial geometry selected might bias the course of the simulation.

Acknowledgements

We wish to thank Phil Crews for assistance in acquiring patellamide C and Chris Ireland for providing a sample of patellamide A. BFM was supported by an EPSRC project grant GR/M37615, and molecular modeling hardware was funded by the Royal Society. LAM was a BBSRC PDRF supported by BBSRC grant number 1/E12737. The Carnegie Trust provided funds for a 6 week visit to Steve Homans' laboratory in Leeds to carry out MD studies. GST was a BBSRC PDRF supported by BBSRC grant number 49/SBD07527. MJ wishes to thank the Aberdeen University Chemistry Department for the provision of new NMR equipment which made this work possible.

References

1 B. S. Davidson, *Chem. Rev.*, 1993, **93**, 1771–1791.
2 T. Ishida, Y. In, F. Shinozaki, M. Doi, D. Yamamoto, Y. Hamada, T. Shioiri, M. Kamigauchi and M. Suguira, *J. Org. Chem.*, 1995, **60**, 3944–3952.
3 M. Doi, F. Shinozaki, Y. In, T. Ishida, D. Yamamoto, M. Kamigauchi, M. Sugiura, Y. Hamada, K. Khoda and T. Shioiri, *Biopolymers*, 1999, **49**, 459–469.
4 D. J. Freeman, G. Pattenden, A. F. Drake and G. Siligardi, *J. Chem. Soc., Perkin Trans. 2*, 1998, 129–135.
5 A. Perczel and M. Hollosi, in *Circular Dichroism and the Conformational Analysis of Biomolecules*; ed. G. D. Fasman, Plenum Press: New York, 1996, p. 738.
6 T. Ishida, M. Tanaka, M. Nabae, M. Inoue, S. Kato, Y. Hamada and T. Shioiri, *J. Org. Chem.*, 1988, **53**, 107–112.
7 L. A. Morris, M. Jaspars, J. J. Kettenes van den Bosch, K. Versluis, A. J. R. Heck, S. M. Kelly and N. C. Price, *Tetrahedron*, 2001, **57**, 3185–3197.
8 A. T. Brünger, X-PLOR, a system for X-ray crystallography and NMR, Version 3.1; Yale University, New Haven, 1993.
9 F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caulfield, G. Chang, T. Hendrickson and W. C. Still, *J. Comput. Chem*, 1990, **11**, 440–467.
10 W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127–6129.
11 R. M. B. Koradi and K. Wuthrich, *J. Mol. Graphics*, 1994, **14**, 51–59.
12 Y. In, M. Doi, M. Inoue, T. Ishida, Y. Hamada and T. Shioiri, *Chem. Pharm. Bull.*, 1993, **41**, 1686–1690.
13 F. J. Schmitz, M. B. Ksebati, J. S. Chang, J. L. Wang, M. B. Hossain, D. van der Helm, M. H. Engel, A. Serban and J. A. Silfer, *J. Org. Chem.*, 1989, **54**, 3463–3472.
14 L. Grondahl, N. Sokolenko, G. Abbenante, D. P. Fairlie, G. R. Hanson and L. R. Gahan, *J. Chem. Soc., Dalton Trans.*, 1999, 1227–1234.

J. Braz. Chem. Soc., Vol. 13, No. 2, 207-217, 2002.
Printed in Brazil - ©2002 Sociedade Brasileira de Química
0103 - 5053 $6.00+0.00

Article

Derivatives of *N,N'*-bis[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]ethanediamide

***John N. Low*[a], *Bruce F. Milne*[a], *Jennifer-Nicola Ross*[a] *and James L. Wardell**[b]**

[a] *Department of Chemistry, University of Aberdeen, Meston Walk, Old Aberdeen AB24 3UE, Scotland*

[b] *Departamento de Química Inorgânica, Instituto de Química, Universidade Federal do Rio de Janeiro, CP 68563, 21945-970, Rio de Janeiro - RJ, Brazil*

O composto $[(HOCH_2)_3CNHC(O)]_2$ (**1**), formado a partir da reação de $(HOCH_2)_3CNH_2$ com EtOC(O)C(O)OEt, reage com aldeídos aromáticos ArCHO, gerando como produtos bis-alquilideno simétricos, *N,N'-bis*(2-Ar-5-$ROCH_2$-1,3-dioxan-5-a)etanodiamides **3** (Ar = Ph, *p*-MeC_6H_4 or *p*-$MeOC_6H_4$, R = H). Uma reação similar com Me_2CO produziu *N'-bis*(2,2-dimethil-5-hidroximetil-1,3-dioxan-5-a)etanodiamida (**2**). Enquanto três estereoisômeros (*Z,Z*)-, (*Z,E*)- e (*E,E*)-**3** (Ar = Ph, R = H) foram formados a partir da reação de **1** com PhCHO, somente (*Z,Z*)-**3** (Ar = *p*-MeC_6H_4 ou *p*-$MeOC_6H_4$, R = H) foi isolado quando **1** reagiu com ArCHO (Ar = *p*-MeC_6H_4 ou *p*-$MeOC_6H_4$). As conformações *Z* têm os grupos: aril- equatorial, $HOCH_2$- equatorial e amido-axial, no sólido e em solução, enquanto as conformações *E* têm os grupos aril- equatorial, amido- equatorial e $HOCH_2$- axial. Uma mistura na proporção 1:1 de (*Z,Z*)-: (*E,E*)-**3** (Ar = Ph, R = H) co-cristaliza. As energias de conformação de (*Z,Z*)- e (*E,E*)-**3** (Ar = Ph, R = H) e **1** foram determinadas por cálculos de mecânica molecular. O estereoisômero (*Z,Z*)-**3** (Ar = Ph, R = H) é mais reativo do que o (*E,E*) em reações de alquilação: somente o estereoisômero (*Z,Z*)-**3** (Ar = Ph, R = Ph_3SnCH_2) foi isolado da reação de uma mistura de (*Z,Z*)- e (*E,E*)-**3** (Ar = Ph, R = H) com Ph_3SnCH_2I. Quando uma mistura 1:1 de (*Z,Z*)- e (*E,E*)-3 (Ar = Ph, R = H) foi reagida com um excesso de brometo alílico, uma mistura na proporção de 4:3 de (*Z,Z*)- e (*E,E*)-**3** (Ar = Ph, R = H_2C=$CHCH_2$) foi isolada. Reação de oxomercuração de (*Z,Z*)- e (*E,E*)-**3** (Ar = Ph, R = H_2C=$CHCH_2$) com $Hg(OAc)_2$ em metanol, seguida por uma troca aniônica utilizando NaCl, produziu um único estereoisômero, {*N*-(*Z*)-[[(R^2)-5-(3-cloromercuria-2-metoxipropil)oximetil]-2-fenil-1,3-dioxan-5-il]}{*N'*-(*Z*)-[[(S^2)-5-(3-cloromercúria-2-metoxipropil)oximetil]-2-fenil-1,3-dioxan-5-il]}etanodiamida (**4**) que foi caracterizado por cristalografia de raio X.

Compound, $[(HOCH_2)_3CNHC(O)]_2$ (**1**), obtained from $(HOCH_2)_3CNH_2$ and EtOC(O)C(O)OEt, reacts with aryl aldehydes, ArCHO, to give the symmetric bis-alkylidene derivatives, *N,N'*-bis(2-Ar-5-$ROCH_2$-1,3-dioxan-5-yl)ethanediamides **3** (Ar = Ph, *p*-MeC_6H_4 or *p*-$MeOC_6H_4$, R = H). A similar reaction with Me_2CO produced *N'*-bis(2,2-dimethyl-5-hydroxymethyl-1,3-dioxan-5-yl)ethanediamide (**2**). While three stereoisomers, (*Z,Z*)-, (*Z,E*)- and (*E,E*)-**3** (Ar = Ph, R = H), were formed from **1** and PhCHO, only (*Z,Z*)-**3** (Ar = *p*-MeC_6H_4 or *p*-$MeOC_6H_4$, R = H) was isolated from ArCHO (Ar = *p*-MeC_6H_4 or *p*-$MeOC_6H_4$) [the *Z* conformations in the solid state and in solution have *equatorial*-aryl, *equatorial*-$HOCH_2$ and *axial*-amido groups: *E* forms have *equatorial*-aryl, *equatorial*-amido and *axial*-$HOCH_2$ groups]. A 1:1 mixture of (*Z,Z*)-:(*E,E*)-**3** (Ar = Ph, R = H) co-crystallises. Molecular mechanics calculations have been carried out on the conformation energies of (*Z,Z*)-and (*E,E*)-**3** (Ar = Ph, R = H) and **1** and support the crystallographic and spectral findings. The stereoisomer, (*Z,Z*)-**3** (Ar = Ph, R = H), is more reactive in alkylation reactions than the (*E,E*)-form: only (*Z,Z*)-**3** (Ar = Ph, R = Ph_3SnCH_2) was isolated from the reaction of a mixture of (*Z,Z*)- and (*E,E*)-**3** (Ar = Ph, R = H) with Ph_3SnCH_2I. From the reaction of excess allyl bromide with a 1:1 mixture of (*Z,Z*)- and (*E,E*)-**3** (Ar = Ph, R = H), a 4:3 mixture of (*Z,Z*)- and (*E,E*)-**3** (Ar = Ph, R = H_2C=$CHCH_2$) was isolated. Oxymercuriation of (*Z,Z*)-**3** (Ar = Ph, R = H_2C=$CHCH_2$) with $Hg(OAc)_2$ in MeOH, followed by anion exchange using NaCl, produced the single stereoisomer, {*N*-(*Z*)-[[(R^2)-5-(3-chloromercuri-2-methoxypropyl)oxymethyl]-2-phenyl-1,3-dioxan-5-yl]}{*N'*-(*Z*)-[[(S^2)-5-(3-chloromercuri-2-methoxypropyl)oxymethyl]-2-phenyl-1,3-dioxan-5-yl]}ethanediamide (**4**), characterised by X-ray crystallography.

Keywords: alkylidene formation, stannylation, oxymercuriation, X-ray crystallography, molecular mechanics

* e-mail: j.wardell@abdn.ac.uk

Introduction

Despite being known for at least half a century,[1] and being readily prepared from common precursors, the symmetric diamido-hexol, *N,N'*-bis[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]ethanediamide, $[(HOCH_2)_3CNHC(O)]_2$ (**1**), has attracted only occasional attention. However, its versatility as a precursor has still been well indicated, *e.g.*, its hexa-*O*-derivatives have been mentioned in patents as potential explosives,[2] as sustained release [nitrogen] fertilizers,[3] in preparations of ink-jets dyes[4] and as antioxidants.[5] Unsubstituted **1** has also been used as a precursor of sucrose mimics[6] and as a ligand in lanthanide complexes in a study of the catalysed hydrolysis of phosphate esters.[7]

The versatility of **1** as a precursor would be greatly enhanced if reactions at the hydroxyl groups were controlled. Such a control effectively requires selective protection of the OH groups. A study of the protection of the OH groups in **1** has been carried out and the findings on the use of alkylidene groups are reported here.

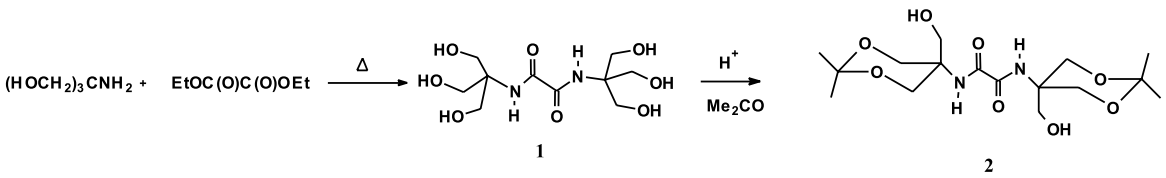
Results and Discussion.

Compound **1** was readily obtained by a published route[6] from $(HOCH_2)_3CNH_2$ and EtO_2CCO_2Et.

*O-Alkylidene derivatives of **1***

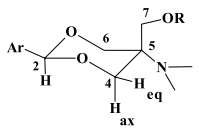
The reaction of a 1:1 mole ratio of **1** with Me_2CO in the presence of a catalytic amount of H_2SO_4 was sluggish and unselective, with both mono-and bis-isopropylidene derivatives being formed. Reaction with excess Me_2CO gave the bis-isopropylidene derivative, *N,N'*-bis(2,2-dimethyl-5-hydroxymethyl-1,3-dioxan-5-yl)ethanediamide (**2**), Scheme 1. Reaction of **1** with excess ArCHO similarly produced *N,N'* -bis-(5-$HOCH_2$-2-Ar-1,3-dioxan-5-yl)ethanediamide **3** (R = H). The chemical shifts for the NH, CH_2OH and 1,3-dioxanyl CH_2 ring protons in **2** are similar to those in (*Z,Z*)-**3** (R = H) and are quite distinct from those in (*E,E*)-**3** (R = H), see Tables 1 and 2 and the later discussion. Thus as in (*Z,Z*)- **3** (R = H)**,** the conformation of the chair shaped 1,3-dioxanyl rings in **2** in solution has the NH(CO) units in axial sites and the CH_2OH (and aryl) groups in equatorial positions. As shown by the NMR spectra, the 1,3-dioxanyl rings are conformationally rigid at ambient temperature with no chair ⇔ chair conversions detected for either the (*Z*) or (*E*)-forms: furthermore no (*Z*)- ⇔ (*E*)- conversions occur.

Three solid stereoisomeric benzylidene products, (*Z,Z*)-, (*Z,E*)- and (*E,E*)-**3** (Ar = Ph, R = H), were isolated from the reaction of **1** with excess PhCHO, see Figure 1. The mole ratios of the initial products, (*Z,Z*)-, (*Z,E*)- and (*E,E*)-**3** (Ar = Ph, R = H), as determined by ^1H NMR spectroscopy, were 1.8: 1.3: 1, *i.e.*, an overall (*Z*): (*E*) ratio of *ca.* 3:2. Recrystallisation of the initial reaction products from aqueous acetone gave as the first crop of crystals, a 1:1 mixture of (*Z,Z*)- and (*E,E*)-**3** (Ar = Ph, R = H), as indicated by NMR spectroscopy and confirmed by X-ray crystallography.[8] Of interest, while co-crystallisation of enantiomers is often reported, reports of the co-crystallisations of other types of isomers, such as the stereoisomers here, are seldom made. This co-crystallized stereoisomeric mixture possessed a wide melting range, 185-224°C. Successive recrystallisations, from ethyl acetate, of the material left in the mother liquor, led to the isolation of crystalline (*Z, Z*)-**3** (Ar = Ph, R = H), m.p. 238-241°C, and subsequently (*Z,E*)-**3** (Ar = Ph, R = H), m.p. 194-197°C. As well as the crystal structure of the 1:1 stereoisomeric mixture, that of the (*Z,Z*)-stereoisomer has also been reported.[8] Suitable crystals of the (*Z,E*)-**3** (Ar = Ph, R = H) isomer could not be grown for crystallography but its stereochemistry was readily resolved from the ^1H and ^{13}C NMR spectra. The (*Z,Z*)-**3** (Ar = Ph, R = H) molecules in the solid state are *Z*-shaped, whereas the (*E,E*)-isomers are maximally extended, as shown in Figure 1. A "*L*"-shape is proposed for the (*Z,E*)-stereoisomer. The (*Z,Z*)-**3** (Ar = Ph, R = H) molecular shapes are essentially identical in the 1:1 stereoisomeric mixture and in the pure single stereoisomer sample. The 1,3-dioxanyl rings adopt chair conformations in all molecules, with the phenyl substituents, at C-2, always in equatorial sites. In (*E*)- and (*Z*)-5-hydroxymethyl-2-phenyl-1,3-dioxane rings, the hydroxymethyl groups at C-5 are in axial and equatorial sites, respectively. There are inversion symmetry related pairs of intramolecular N-H—O hydrogen bonds between amide N-H and adjacent carbonyl

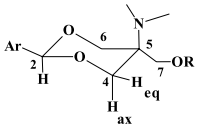


Scheme 1.

Table 1. ^{1}H NMR data for **2** and **3** in DMSO-d_6



E form



Z form

Compound	H-2 (s)	H-4(H-6) (d) ax eq; [J(H_{ax}-H_{eq})]	NH (s)	H-7 [J(H,OH)]	OH [J(H,OH)]	Phenyl [J(H-H)]	Others
2	-	3.75 4.13 [11.7]	7.93	3.68 [5.8]	5.04(t) [5.8]	-	1.31 & 1.34 (Me)
(Z,Z)-**3** (Ar = Ph, R = H)	5.55	3.96 4.47 [11.6]	8.02	3.64 [5.7]	5.03(t) [5.7]	7.34-7.43(m)	
(Z,E)-**3** (Ar = Ph, R = H)	5.54	3.96 4.46 [11.6]	7.99	3.63 [5.75]	5.00(t) [5.7]	7.34-7.43(m)	
	5.56	4.22 [0]	8.11	3.93 [5.75]	5.21(t) [5.8]		
(E,E)-**3** (Ar = Ph, R = H)	5.57	4.23 4.24 [10.9]	8.10	3.96 [5.8]	5.20(t) [5.8]	7.34-7.43(m)	
(Z,Z)-**3** (Ar = *p*-$MeOC_6H_4$ R = H)	5.50	3.81 4.45 [11.6]	8.02	3.65 [2.8]	5.02(t) [2.8]	6.91(d)& 7.34(d) [8.7]	3.74 (OMe)
(Z,Z)-**3** (Ar = *p*-MeC_6H_4 R = H)	5.51	3.96 4.46 [11.6]	8.02	3.65 [-]	4.25-4.75(br) [-]	7.15(d)& 7.24(d) [8.0]	2.28 (Me)
(Z,Z)-**3** (Ar = Ph, R = $MeSO_2$)	5.62-5.67	4.01-4.23	8.23	4.50 [-]	-	7.35-7.57 (m)	2.38 (Me)
(E,E)-**3** (Ar = Ph, R = $MeSO_2$)	5.62-5.67	4.41-4.59	8.44	4.83 [-]	-	7.35-7.57 (m)	2.50 (Me)
(Z,Z)-**3** (Ar = Ph, R = H_2C=$CHCH_2$)	5.59	3.99 4.53 [11.5]	8.18	3.65 (s)	-	7.34-7.39(m)	3.94 5.77-5.90, 5.26, 5.16
(E,E)-**3** (Ar = Ph, R = H_2C=$CHCH_2$)	5.57	4.10 4.35 [11.2]	8.27	3.93-4.08 (m)	-	7.34-7.46(m)	3.93 5.78-5.98, 5.13-5.34
(Z,Z)-**3** (Ar = Ph, R = Ph_3SnCH_2)[a]	5.28	3.88 4.30 [11.8]	7.87	4.31	-	7.37-7.50 7.59-7.64	3.82 ($SnCH_2$)

[a] in $CDCl_3$

Table 2. ^{13}C NMR and ^{119}Sn NMR data for **2** and **3** in DMSO-d_6

Compound	C-2	C-4 (C-6)	C5	C(O)	C-7	Others, δ^{13}C, unless stated
2	99.3	62.6	57.7	161.2	61.5	24.6 & 25.7 (Me)
(*Z,Z*)-**3** (Ar = Ph, R = H)	102.1	70.6	56.4	161.3	61.0	127.7, 129.7, 130.4 & 139.7 (Ph)
(*E,E*)-**3** (Ar = Ph, R = H)	102.4	69.4	54.5	161.3	61.3	128.0, 129.7, 130.4 & 139.6 (Ph)
(*Z,E*)-**3** (Ar = Ph, R = H)	102.1 102.4	70.6 69.4	56.4 54.5	161.1 161.5	61.0 61.3	127.7, 127.8,129.6, 130.5 & 139.6, 139.7 (Ph)
(*Z,Z*)-**3** (Ar = *p*-$MeOC_6H_4$ R = H)	102.1	70.6	56.3	161.3	61.0	56.7 (OMe), 115.0, 129.1, 132.1 & 161.1 (Ph)
(*Z,Z*)-**3** (Ar = Ph, R = $MeSO_2$)	102.4	69.9	54.5	161.7	53.7	38.6, 127.8, 129.6, 130.6 & 139.4 (Ph)
(*E,E*)-**3** (Ar = Ph, R = $MeSO_2$)	102.9	69.7	53.3	161.6	54.9	38.7, 127.8, 129.7, 130.6 &139.3 (Ph)
(*Z,Z*)-**3** (Ar = Ph, R = H_2C=$CHCH_2$)	102.4	69.5	55.6	161.4	70.8	73.4, 118.4 & 136.4 (allyl), 127.7, 129.6, 130.4 & 139.4 (Ph)
(*E,E*)-**3** (Ar = Ph, R = H_2C=$CHCH_2$)	102.2	69.1	53.9	161.3	70.1	73.1, 118.2 & 136.3 (allyl), 127.8, 129.8, 130.5 & 139.6 (Ph)
(*Z,Z*)-**3** (Ar = Ph, R = Ph_3SnCH_2) in $CDCl_3$	102.6	70.6	54.3	160.5	74.6[a]	64.0[b], 126.1, 128.2, 128.6 & 129.1 (PhCH) 136.7[c], 137.0[d], 137.2[e] & 137.6[f] (PhSn) δ^{119}Sn -139.8

[a] J(^{119,117}Sn-^{13}C) = 54Hz; [b] J(^{119,117}Sn-^{13}C) = 470, 452Hz; [c] J(^{119,117}Sn-^{13}C-p) = 17Hz; [d] J(^{119,117}Sn-^{13}C-m) = 44Hz; [e] J(^{119,117}Sn-^{13}C-o) = 33Hz; [f] J(^{119,117}Sn-^{13}C) = 486, 470Hz

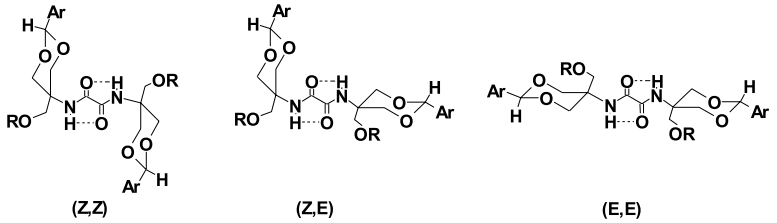


Figure 1. Stereoisomers of *N,N'*-bis(2-aryl-5-alkoxymethyl-1,3-dioxan-5-yl)ethanediamide (**3**)

oxygen centres: intermolecular O-H—O hydrogen bonding between hydroxymethyl group and carbonyl atoms link molecules.[8]

From reaction of **1** with *p*-methoxybenzaldehyde or *p*-methylbenzaldehyde, only the (*Z,Z*)-isomer of **3** (Ar = *p*-R'C_6H_4, R' = MeO or Me, R = H) was isolated. NMR spectra clearly indicated their stereochemistries to be (*Z,Z*), with no other isomer present in the isolated and toluene-washed compounds. Both the (*Z,Z*)-isomers were essentially insoluble in most common organic solvents, and only sparingly so in DMSO. They were obtained from the reaction mixtures as pure compounds merely by extensive extractions with toluene, to remove the excess aldehyde. Small quantities of the other stereoisomers could have been lost with the washings: as there was no indications in the NMR spectra for other stereoisomers in the original crude product mixtures, the maximum amounts of these would have to be less than 5%.

*Molecular mechanics calculations on **1** and **3** (Ar = Ph, R = H)*

Molecular mechanics calculations on **1** and **3** (Ar = Ph, R = H) were obtained[9] using the program Macromodel v6.5. Molecules in the gas phase were investigated using a 5000 step Monte Carlo search for conformers, followed by energy minimisation of the conformers generated. All energy minimisations were performed with the Macromodel MM2* force field: solvent effects were assessed by the GB/AS continuum solvent model.[10]

Calculations on **1** and **3** (Ar = Ph, R = H) were limited to symmetric structures, *e.g.*, only the (*Z,Z*)- and (*E,E*)- and not (*Z,E*)-forms of **3** were included in the calculations. Two favoured conformations, **1a** and **1b**, were calculated for **1**, the more stable form in the three phases considered being invariably **1a**, see Figure 2 and Table 3**.** In both conformations, one of the CH_2OH groups in each $C(CH_2OH)_3$ unit is H-bonded to the carbonyl oxygen of the adjacent amide. The distinction between **1a** and **1b** resides in the orientations of the other two CH_2OH units, see Figure 2: these two OH groups are considered to be

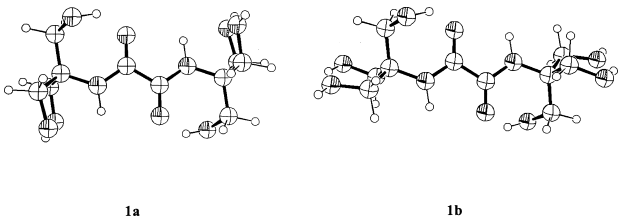


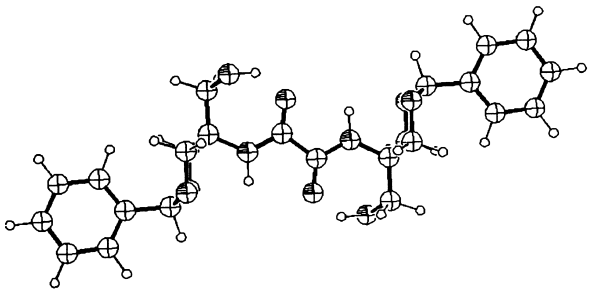
Figure 2. Symmetric conformations of **1**

Table 3. Molecular mechanics (MM2*) calculated energies for conformations of **1** and **3**

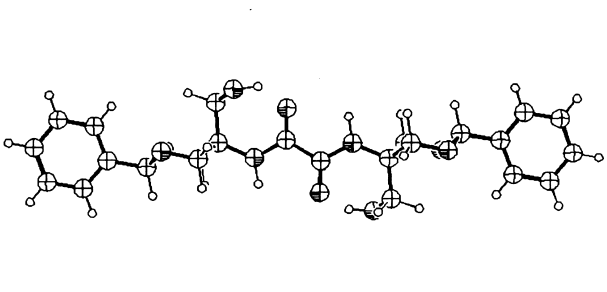
Conformation	Energy ($KJmol^{-1}$)		
	gas phase	***H_2O***	***$CHCl_3$***
1a	-113.2	-187.46	-161.98
1b	-93.04	-173.19	-145.95
[(*E,E*)-**3a**]	-357.07	-384.87	-428.95
[(*E,E*)-**3b**]	-336.26	-373.54	-409.03
[(*Z,Z*)-**3b**]	-362.32	-371.94	-404.44
[(*Z,Z*)-**3a**]	-362.32	-397.37	-438.07

those used in the formation of the 1,3-dioxanyl rings in **3** (Ar = Ph, R = H). Calculations were carried out on four conformers of **3** (Ar = Ph, R = H) - two (*Z,Z*)- and two (*E,E*)-forms in three different phases, see Figure 3. Differences in the relative energies are found between the three phases, indicating the importance of solvation and H-bonding. The relative energies calculated for the organic solvent, $CHCl_3$, are most relevant to the reactions carried out in the aryl aldehyde.

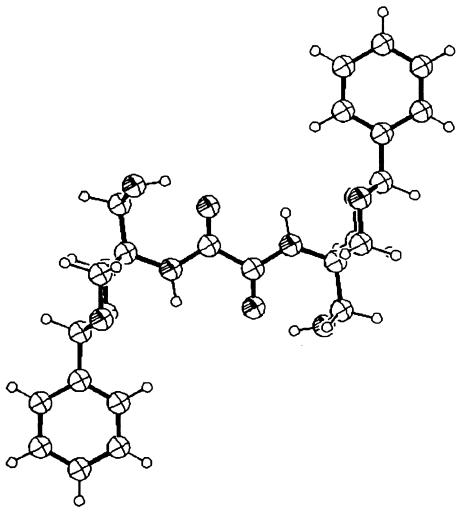
Reaction of PhCHO with the more stable **1a** conformer provides either the (*E,E*)-conformer [(*E,E*)-**3a**] (Ph axial/ CH_2OH equatorial) or the (*Z,Z*)- conformer [(*Z,Z*)-**3a**] (Ph equatorial/ CH_2OH equatorial): calculations show that [(*Z,Z*)-**3a**] is the more stable in $CHCl_3$. Reaction of PhCHO with the other conformer **1b** will provide either the (*E,E*) conformer [(*E,E*)-**3b**] (Ph equatorial/ CH_2OH axial) or the (*Z,Z*)-conformer [(*Z,Z*)-**3b**] (Ph axial/ CH_2OH axial). Calculations indicated that of these two, [(*E,E*)-**3b**] is favoured. Thus the molecular mechanics calculations point to the favoured formations of [(*Z,Z*)-**3a**] and [(*E,E*)-**3b**] with the former dominating, from the two symmetrical forms of



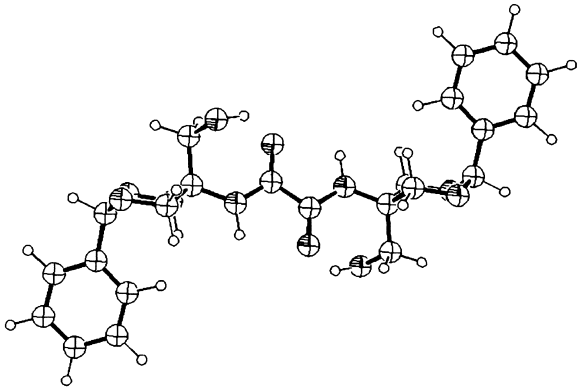
[(*E,E)*-**3a**]



[(*E,E)*-**3b**]



[(*Z,Z)*-**3a**]



[(*Z,Z*)-**3b**]

Figure 3. Symmetric conformers of **3**

1. The calculated preference for the *Z*-conformations of the 1,3-dioxanyl rings [*i.e.* with axial amido groups] and the indications that the phenyl groups are generally in equatorial sites are in agreement with the NMR and crystallographic findings.

The preference for the amido group over the CH_2OH group to occupy axial sites in **3** is similar to the situation reported for cyclohexane derivatives. In cyclohexanes, substituents, X, invariably favour equatorial sites. For example, the preferences [conformational energies][11] for NHC(O)OPh, CH_2OH and Ph to be in equatorial sites in cyclohexanes are 6.7, 7.36 and 11.7 kJ mol^{-1}, respectively, the NHC(O)Ph group being the nearest to the NHCOCONH group for which data could be found.[11] However, this equatorial preference need not apply to X-substituted 1,3-dioxanes, where polar O—X interactions have to be considered. As also pointed out by Eliel and Wilen,[11] steric compression of axial groups by lone pairs on oxygen atoms in 1,3-dioxanes is very small compared to the compression by syn-axial H-atoms in cyclohexanes. These effects can, in fact, result in a preference for axial positions, as shown, for example, by the CH_2OH group in 5-hydroxymethyl-2-isopropyl-1,3-dioxane in CCl_4 solution.[11]

Reactions of the alkylidene derivatives

The remaining free hydroxyl groups in **3** (Ar = Ph, R = H) can be successfully derivatised as illustrated by the use of methanesulfonyl chloride and allyl bromide. Reactions of the 1:1 (*Z,Z*): (*E,E*)-**3** mixture with these reagents gave 1:1 (*Z,Z*): (*E,E*)-mixtures of *N,N'*-bis(5-mesyloxymethyl-2-phenyl-1,3-dioxan-5-yl)ethanediamides **3** (Ar = Ph, R = $MeSO_2$), and *N,N'*-bis(5-allyloxymethyl-2-phenyl-1,3-dioxan-5-yl)ethanediamides **3** (Ar = Ph, R = CH_2=$CHCH_2$), respectively, as shown by the NMR spectra. Work-up can lead to preferential isolation of one of the components, *e.g.*, [(*Z,Z*)-**3** (Ar = Ph, R = H_2C=$CHCH_2$)] was isolated in a pure form on recrystallisation of the 1:1 mixture of

[(*Z,Z*):(*E,E*)-**3** (Ar = Ph, R = CH_2=$CHCH_2$)], from ethyl acetate: further work-up of the mother liquor, unfortunately, resulted in irreversible changes, which prevented the collection of [(*E,E*)-**3** (Ar = Ph, R = CH_2=$CHCH_2$)]. In general, separation of both sets of isomers could be achieved by chromatography. Attempts to obtain *N,N'*-bis(5-chloromethyl-2-phenyl-1,3-dioxan-5-yl)ethanediamides from **3** (Ar = Ph, R = H) using Ph_3P/CCl_4 or thionyl chloride were unsuccessful: no reaction occurred using Ph_3P/CCl_4, even on reflux, while extensive decomposition resulted from the use of thionyl chloride.

Organometallic derivatives

Bis-oxymercuriation of (*Z,Z*)-**3** (Ar = Ph, R = CH_2=$CHCH_2$) using $Hg(OAc)_2$ in methanol proceeded readily at room temperature.[12] NMR spectroscopy indicated that the product, isolated on crystallisation from Me_2CO/ethyl acetate, after anion exchange with NaCl, was the Marknovikov adduct, *N,N'*-bis-{(*Z*)-[5-[(3-chloromercuri-2-methoxypropyl)oxymethyl]-2-phenyl-1,3-dioxan-5-yl]}ethanediamide (**4**) Equation 1. The atom arrangements in **4** were confirmed by X-ray crystallography. Furthermore, the stereochemistries at the -CH(OMe)- centres [C(19) and C(19a) in the crystallographic numbering scheme] were revealed, see Figure 4. Unfortunately, the structure of **4** was only refined to 9.4%, due in the greater part to crystal decomposition in the X-ray beam.

The molecule of **4** possesses a centre of symmetry at the midpoint of the central (O)-C-C(O) bond with both mercury sites being equivalent. Due to the symmetry in **4**, selected geometric parameters for only one half of the molecule are listed in Table 4: parameters involving atoms in the other half of the molecule are indicated by the use of the superscipt ([a]) [symmetry operation *a*: -x, 2-y, 1-z]. The carbon atoms at which the OMe groups are attached have opposite chirality and so overall the molecule is achiral, *i.e.* we have {*N*-(*Z*)-[[(R^2)-5-(3-chloromercuri-2-methoxypropyl)oxymethyl]-2-phenyl-1,3-dioxan-5-yl]}{*N'*-(*Z*)-[[(S^2)-5-(3-chloromercuri-2-methoxypropyl)oxymethyl]-2-phenyl-1,3-dioxan-5-yl]}ethanediamide stereoisomer of **4**: the high yield of the isolated stereoisomer pointed to a

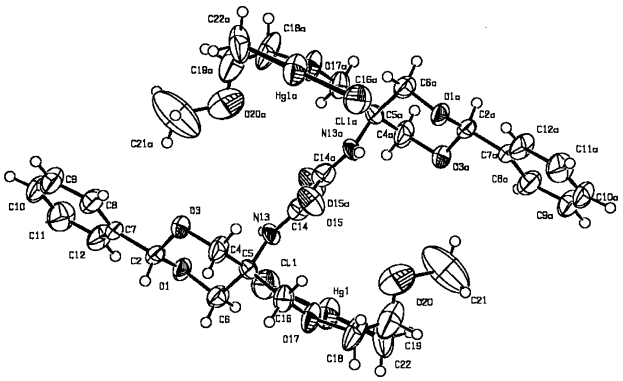
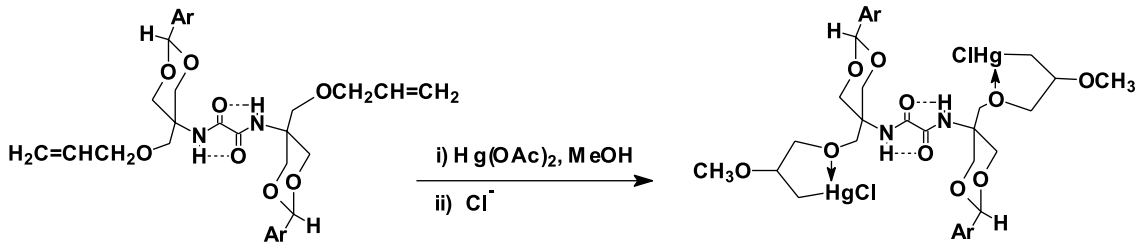


Figure 4. Atom numbering system and atom arrangements in **4.** Probability ellipsoids drawn at the 50% level



(1)

Table 4. Selected geometric parameters (Å,°) for **4**

Hg(1)-Cl(1)	2.331(6)	Hg(1)-C(22)	2.06(3)
Hg(1)-O(17)	2.882(15)	Hg(1)-O(20)	2.04(3)
O(15)-C(14)	1.29(2)	C(14)-C(14a)	1.520(5)
N(13)-C(5)	1.48(2)	N(13)-C(14)	1.22(2)
Cl(1)-Hg(1)-C(22)	179.6(7)	O(15)-C(14)-N(13)	129.2(17)
C(22)-Hg-O(17)	71.0(9)	C(22)-Hg-O(20)	43.8(8)
Cl(1)-Hg-O(17)	109.5(3)	Cl(1)-Hg-O(20)	136.2(5)
O(17)-Hg-O(20)	53.9(7)	O(15)-C(14)-C(14)	111.5(15)
N(13)-C(14)-C(14)[a]	118.9(16)	C(5)-N(13)-C(14)	125.4(14)
Hg(1)—Hg(1)[i]	3.775(2)		
O(20)-C(19)-C(22)-Hg(1)	-63(3)°	C(18)-C(19)-C(22)-Hg(1)	62(3)°

a: symmetry operation: -x, 2-y, 1-z; *i*: symmetry operation: -x, 3-y, -z

high stereoselectivity in the mercuriation reaction. Indeed other stereoisomers of **4**, *e.g.* the bis-[(*N,N'*-(*Z*),(*R*^{2})] or bis-[(*N,N'*-(*Z*),(*S*^{2})] stereoisomers, were not detected. The dioxanyl rings, as found for stereoisomers of **3** (Ar = Ph: R = H) have chair conformations, with puckering parameters,[13] amplitude (Q) = 0.570(16)°, θ = 175.3(17)°, ϕ = 191(25)°.

The primary bonds to each Hg in **4** are essentially collinear, C(22)-Hg(1)-Cl(1) / [C(22)^a-Hg(1)^a-Cl(1)^a] = 179.6(7)°, see Table 4: the primary bond lengths, Hg(1)-C(22) / [Hg(1)^a-C(22)^a] = 2.06(3)Å and Hg(1)-Cl(1) / [Hg(1)^a-Cl(1)^a] = 2.331(6)Å, are in the expected regions.[14-16] As well as forming primary bonds, mercury(II) also exhibits a strong tendency to form secondary bonds within the sum of the van der Waals radii of the relevant atoms.[14-17] There is some controversy over the value of the van der Waals radius of mercury. Values of 1.73–2.00Å have been suggested by Canty and Deacon,[18] while even higher values, 2.1-2.2Å, have been proposed by Batsanov,[19] all these being much higher than the value of 1.55Å, estimated from the critical volume of the metal.[20] As the generally accepted van der Waals radius for O is 1.50Å, the O(17)-Hg(1) /[O(17)a—Hg(1)a] intramolecular separations of 2.882(15)Å fall well within the van der Waals radii sum for Hg and O, no matter which Hg vdW value is taken. These secondary Hg-O bonds result in the formation of 5-membered Hg(1)-O(17)-C(18)-C(19)-C(22) / [Hg(1)^a-O(17)^a-C(18)^a-C(19)^a-C(22)^a] rings with envelope conformations [flap at C(19) / C(19)^a]. Other intramolecular, Hg—O separations, Hg(1)-O(20) / [Hg(1)^a-O(20)^a] = 3.32(3)Å, are within the majority of the limits set for the sum of the van der Waals radii for Hg and O. As shown in Figure 5, the gauche arrangements of Hg(1) and O(20), Hg(1) and C(18), about the C(19)-C(22) bond, are ideal for Hg—O (20) and Hg—O(17) interactions and are taken as indications of secondary bonding.

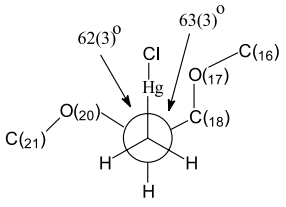


Figure 5. Arrangement at the mercury centre in **4** looking down the C(22)-C(19) bond

There are, in addition, intermolecular contacts, Hg—Hg^i [3.776(2)Å] in **4**: symmetry operation *i*: -x, 3-y, -z], which produce staggered chains of molecules almost parallel to the *b*, *c* diagonal. Intermolecular mercury—mercury interactions have been variously reported, for example,[21] Hg—Hg = 3.5620(5)Å in [Hg$_3$O(1,3-dimethyl-uracil-5-yl)$_3$]NO$_3$.2H$_2$O. The closest Hg separations with aromatic C atoms in **4** occur at distances of 3.75-3.79Å, *cf.* vdW radius of C = 1.65-1.70Å.[20] As found generally in organomercury chemistry, secondary bonds have little effect on the linear geometry of the primary bonds.[14-16]

As with **3** (Ar = Ph, R = H), there are intramolecular hydrogen bonds involving N(13)-H(13)—O(15) [N(13)-H(13) = 0.86Å: H(13)—O(15) = 2.28Å: N(13)—O(15) 2.67(2) Å: N(13)-H(13)—O(15) = 108°] and analogously for N(13)^a-H(13)^a—O(15)^a.

Attempts to get Ph$_3$SnH addition to the alkenyl groups in (*Z,Z*)-**3** (Ar = Ph, R = CH$_2$=CHCH$_2$), in the presence of the radical initiator, AIBN, failed: complete recover of [(*Z,Z*)-**3** (Ar = Ph, R = CH$_2$=CHCH$_2$)] was made and a near quantitative formation of Ph$_3$SnSnPh$_3$ was indicated. Tin hydrides, such as Ph$_3$SnH, are known to be sensitive to organic bases, such as amines and sulfoxides, and decompose to Ph$_3$SnSnPh$_3$:[22] it appears that the amido derivative, (*Z,Z*)-**3** (Ar = Ph, R = CH$_2$=CHCH$_2$), is also sufficiently basic to affect this decomposition.

A bis-stannylated derivative was, however, obtained from **3** (Ar = Ph, R = H): the compound, (*Z,Z*)-**3** (Ar = Ph, R = Ph$_3$SnCH$_2$), d^{119}Sn = -139.8ppm, was isolated from the reaction of the 1:1 mixture of (*Z,Z*)-:(*E,E*)-**3** (Ar = Ph, R = H), with Ph$_3$SnCH$_2$I [1:2.05 mole ratio], in DMF in the presence of NaH, after chromatography, see Figure 6. The ^{119}Sn NMR spectrum of the crude reaction product, prior to chromatographic separation, exhibited several peaks: the most intense peak being at -139.8ppm, *i.e.* the precise value of the isolated product. Other peaks were also in the region expected for Ph$_3$SnCH$_2$OR* compounds [-145 to -135ppm],[23] but all these were minor peaks. It is possible

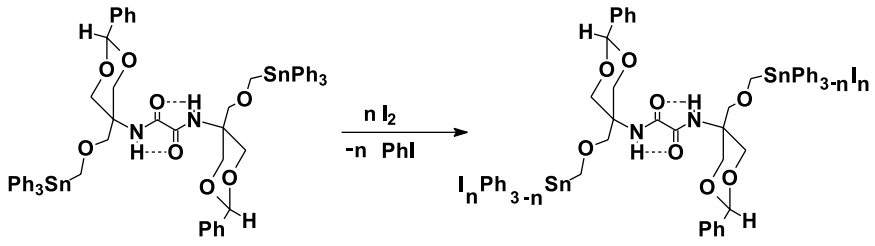


Figure 6. (*Z,Z*)-**3** (Ar = Ph, R = Ph$_3$SnCH$_2$) and reactions with iodine

that the stannylated derivative of (*E*,*E*)-**3** (Ar = Ph, R = H) would also have a δ^{119}Sn value of -139.8ppm, and that it was also present among the reaction products, but it somehow escaped isolation by chromatography. However, the latter is considered unlikely, and it is assumed that (*E*,*E*)-**3** (Ar = Ph, R = Ph_3SnCH_2) could not be a significant product, if formed at all. The crystal structure determination of (*Z*,*Z*)-**3** (Ar = Ph, R = Ph_3SnCH_2), already reported,[24] indicated the tin centres to be 4-coordinate, with slightly distorted tetrahedral geometries, as have been found for other Ph_3SnOR* compounds.[23] As also in these other Ph_3SnCH_2OR* compounds, the intramolecular Sn—O(C-5) separation is short [2.91(1)Å] and within the limits accepted for coordination [the sum of the van der Waals radii for Sn and O is *ca.* 4.1Å].[20] However, no Sn-O bonding is assumed since the C-Sn-C angles are close to those expected for a 4-coordinate and near tetrahedral arrangement. Any Sn—O interaction would create, in any case, highly strained three membered rings.

As shown by the structures listed in the Cambridge Crystallographic Data Base,[25] tetraorganotin compounds generally have 4-coordinate tin centres with near tetrahedral geometries. This is a consequence of the poor acceptor strength of the tin centre: however, in a limited number of rigid tetraorganotin compounds with suitably sited donor groups, intramolecular complexation can occur with the formation of 5- and even 6-coordinate tin centres.[26] Clearly the tin centres in (*Z*,*Z*)-**3** (Ar = Ph, R = Ph_3SnCH_2) are not complexed to any of the donor centres. The Lewis acidity of organotin halides, $R_{4-n}SnX_n$, invariably increases with the value of n. The reactions of (*Z*,*Z*)-**3** (Ar = Ph, R = Ph_3SnCH_2) with iodine (2n mol equivalents) were undertaken to obtain the stronger Lewis acids, (*Z*,*Z*)-**3** (Ar = Ph, R = $I_nPh_{3-n}SnCH_2$, n = 1 or 2): as indicated by the solution NMR parameters, especially δ^{119}Sn values, (*Z*,*Z*)-**3** (Ar = Ph, R = IPh_2SnCH_2) [δ^{119}Sn = -122.5ppm] and (*Z*,*Z*)-**3** (Ar = Ph, R = $I_2PhSnCH_2$) [δ^{119}Sn = -227ppm], were 4 and 5 coordinate, respectively.

Conclusions

Protection of 4 of the 6 hydroxyl groups in **2** can be achieved by standard alkylidene procedures: the bis alkylidene derivatives, **3**, can be further derivatised by standard means. Transformations of the bis-mercurated, **4**, and the bis stannylated, (*Z*,*Z*)-**3** (Ar = Ph, R = Ph_3SnCH_2), compounds provide further routes to selected derivatives.

Experimental

Melting points were determined using a Kofler hotstage and are uncorrected Solution NMR spectra were obtained on Bruker 250 MHz and Varian 400 MHz instruments. IR spectra were obtained on Philips Analytical PU 9800 FTIR and Nicolet 205 FTIR instruments.

*N,N' -Bis-[tris(hydroxymethyl)methyl]ethanediamide (**1**)*

A solution of tris(hydroxymethyl)aminomethane (10.00 g, 82.6 mmol) and diethyl oxalate (6.03g, 41.3 mmol) in methanol (25ml) was refluxed for 2h. The reaction mixture was cooled, filtered and the precipitate was recrystallised from aqueous EtOH (80%); yield 75-80%; m.p. 221-223 °C [lit.[6]m.p. 224 °C]. ^{1}H and ^{13}C NMR spectra were identical with those reported.[6] IR (ν_{max}/cm^{-1}): 3413, 3339 (NH), 3245 (OH), 1667, 1516 (amide),

N,N' -Bis-(5-hydroxymethyl-2,2-dimethyl-1,3-dioxan-5-yl)ethanediamide (**2**)

A reaction mixture of **1** (3.00g, 10mmol) and concentrated H_2SO_4 (1ml) in acetone (100ml) was stirred at room temperature for 16h and filtered. The filtrate was washed with aqueous sodium bicarbonate, dried and rotary evaporated. The solid residue was recrystallised from acetone, yield 15%, m. p. 157-159 °C. Anal. calcd. for $C_{16}H_{28}N_2O_8$: C, 51.1; H, 7.5; N, 7.4. Found: C, 51.0; H, 7.5; N, 7.4%. NMR data are listed in Table 1. IR (ν_{max}/cm^{-1}): 3312, 3211, 3147, 3048, 1663, 1574 (amide).

*N,N' -Bis-(5-hydroxymethyl-2-phenyl-1,3-dioxan-5-yl)ethanediamide, **3** (Ar = Ph, R = H)*

A reaction mixture of **1** (1.5g, 5.1mmol) and concentrated H_2SO_4 (0.5ml) in benzaldehyde (25ml) was stirred for 12h at room temperature. Diethyl ether (10ml) and water (25ml) were added, and the viscous organic layer separated. On standing, a finely divided solid was collected, yield 85-95%. NMR spectra of this crude product revealed the presence of three stereoisomers, (*Z*,*Z*)-, (*E*,*E*)- and (*Z*,*E*)-**3** (Ar = Ph, R = H), in mole ratios of 1.8:1.3:1. Recrystallisation of the crude product from aqueous acetone gave a 1:1 mixture of stereoisomers, (*Z*,*Z*)- and (*E*,*E*)-**3** (Ar = Ph, R = H), yields ranged between 45-60%. This 1:1 mixture had a melting range between 185-225 °C. Anal. calcd. for $C_{24}H_{28}N_2O_8$: C, 61.0; H, 6.0; N, 5.9. Found: C, 61.1; H, 6.0; N, 5.9 %. IR (ν_{max}/cm^{-1}): 3490, 3382, 3337, 1671, 1516.

All solvents were removed from the mother liquor under vacuo to leave a solid residue, which was recrystallised from ethyl acetate. The colourless crystalline solid obtained was the single steroisomer, (*Z*,*Z*)-**3** (Ar = Ph, R = H); m.p. 238-241 °C. The mother liquor from the second crystallisation, on concentration and further recrystallisation gave

crystals of stereoisomer, (*Z,E*)-**3** (Ar = Ph, R = H); m.p. 194-197 °C. NMR spectra of the three stereoisomers are displayed in Tables 1 and 2.

*N,N' -Bis-[(Z)-(5-hydroxymethyl-2-(p-methoxyphenyl)-1,3-dioxan-5-yl)]ethanediamide, (Z,Z)-**3** (Ar = p-MeOC$_6$H$_4$, R = H)*

A reaction mixture of **1** (1.5g, 5.1mmol) and concentrated H_2SO_4 (0.5ml) in *p*-methoxybenzaldehyde (25ml) was stirred for 12h at room temperature. Diethyl ether (10ml) and water (25ml) were added, and the viscous organic layer separated. On standing, a finely divided solid was collected. ^{1}H NMR spectrum of the solid, indicated the presence of a single stereoisomer, still contaminated with *p*-methoxybenzaldehyde. The latter was removed on Soxhlet extraction with toluene. The white residue was insoluble in most organic solvents and only sparingly soluble in DMSO. Yield: 78%; m.p. 245-248 °C. Anal. calcd. for $C_{26}H_{32}N_2O_{10}$: C, 58.6; H, 6.1; N, 5.3. Found: C, 58.1; H, 5.9; N, 5.6 %. IR (ν_{max}/cm^{-1}): 3467, 3412, 3374, 3341, 1676, 1518. NMR data, displayed in Tables 1 and 2, indicated that the stereoisomer was (*Z,Z*)-(**3**: Ar = *p*-MeOC$_6$H$_4$, R = H).

*N,N' -Bis-[(Z)-(5-hydroxymethyl-2-(p-methylphenyl)-1,3-dioxan-5-yl)]ethanediamide, (Z,Z)-**3** (Ar = p-MeC$_6$H$_4$, R = H)*

A reaction mixture of **1** (1.5g, 5.1mmol) and concentrated H_2SO_4 (0.5ml) in *p*-methylbenzaldehyde (25ml) was stirred for 12h at room temperature. Diethyl ether (10ml) and water (25ml) were added, and the viscous organic layer separated. On standing, a finely divided solid was collected. ^{1}H NMR spectrum of the solid, indicated the presence of a single stereoisomer, (*Z,Z*)-**3** (Ar = *p*-MeC$_6$H$_4$, R = H), still contaminated with *p*-methylbenzaldehyde, which was removed on Soxhlet extraction with toluene. The white residue was insoluble in most organic solvents and only sparingly soluble in DMSO. Yield: 73%; m.p. 240 °C dec. Anal. calcd. for $C_{26}H_{32}N_2O_8$: C, 62.4; H, 6.4; N, 5.6. Found: C, 61.7; H, 6.3; N, 5.6 %. IR (ν_{max}/cm^{-1}): 3488, 3366, 3336, 1686, 1508.

NMR data are displayed in Tables 1 and 2.

*N,N' -Bis-(Z)- and -(E)- (5-allyloxymethyl-2-phenyl-1,3-dioxan-5-yl)ethanediamide, (Z,Z)- and (E,E)-**3** (Ar = Ph, R = H$_2$C=CHCH$_2$)*

To a solution containing a 1:1 mixture of (*Z,Z*)-:(*E,E*)-**3** (Ar = Ph, R = H), (2.00g, 4.23mmol) and NaH (0.2g) in dry DMF (6ml) was added allyl bromide (3.07g, 25.4 mmol). After stirring the reaction mixture at room temperature for 24h, it was poured onto ice (25g) and filtered. The dried solid residue had a melting range from 137 to 184 °C, and was shown by the ^{1}H NMR spectrum to be a 4:3 mixture of stereoisomers, (*Z,Z*)-:(*E,E*)-**3** (Ar = Ph, R = H$_2$C=CHCH$_2$. Recrystallisation from acetone gave pure stereoisomer (*Z,Z*)-**3** (Ar = Ph, R = H$_2$C=CHCH$_2$), m.p. 155-156 °C. The mother liquor still contained both (*Z,Z*)-:(*E,E*)-**3** (Ar = Ph, R = H$_2$C=CHCH$_2$). NMR data are displayed in Tables 1 and 2.

*N,N'-Bis-{(Z)-[5-(3-chloromercuri-2-methoxypropyl-oxymethyl)-2-phenyl-1,3-dioxan-5-yl]}ethanediamide (**4**)*

To a suspension of (*Z,Z*)-**3** (Ar = Ph, R = H$_2$C=CHCH$_2$) (0.38g, 1.2 mmol) in MeOH (10ml) was added a solution of Hg(OAc)$_2$ (0.44g, 1.2mmol) in MeOH (20ml). After stirring overnight, a saturated solution of sodium chloride (20 ml) was added, followed by water (80 ml). The precipitate was collected, and recrystallised from acetone: ethyl acetate (2:1) to give colourless crystals of {*N*-(*Z*)-[[(*R*^{2})-5-(3-chloromercuri-2-methoxypropyl)oxymethyl]-2-phenyl-1,3-dioxan-5-yl]}{*N*'-(*Z*)-[[(*S*^{2})-5-(3-chloromercuri-2-methoxypropyl)oxymethyl]-2-phenyl-1,3-dioxan-5-yl]}ethanediamide (**4**). Yield 95%, m.p.169-171 °C. ^{1}H NMR (CDCl$_3$): δ: 1.84-1.97 (m, 4H, CH$_2$Hg), 3.31(s, 6H, OMe), 3.46-3.50 (m, 4H, ≡COCH$_2$), 3.74-3.82 (m, 4H, C*H*$_2$OCHPh), 3.94-4.03 (m, 6H, =C*H*OCH$_3$ + OC*H*$_2$CHOCH$_3$), 4.49-5.58 (m, 4H, C*H*$_2$OCHPh), 5.61 (s, 2H, C*H*Ph), 7.36-7.39 (m, 6H, aryl), 7.49-7.52 (m, 4H, aryl), 7.98 (s, 2H, NH). ^{13}C NMR (CDCl$_3$): δ: 31.8 (CH$_2$Hg), 53.9 (C$_{quart}$), 56.4 (OMe), 69.6 (≡C-*C*H$_2$O), 69.9 (CH$_2$), 70.3 (=*C*HOCH$_3$), 73.7 (*C*H$_2$CHOCH$_3$), 101.5(*C*HPh), 126.1, 128.2, 129.0 & 137.3 (aryl), 159.7 (CO).

*N,N'-Bis-[(Z)-(5-triphenylstannylmethyloxymethyl)-2-phenyl-1,3-dioxan-5-yl)]ethane-diamide, (Z,Z)-**3** (Ar = Ph, R = Ph$_3$SnCH$_2$)*

Sodium hydride (1.50g, 6.5 mmol), and (iodomethyl) triphenylstannane (4.28g, 8.72 mmol) were successively added to a solution of a 1:1 mixture of (*Z,Z*)-:(*E,E*)-**3** (Ar = Ph, R = H), (2.00g, 4.23 mmol) in anhydrous DMF (20ml). The mixture was stirred overnight at room temperature, water (20ml) and diethyl ether (20ml) were added, and the organic layer was collected, dried over magnesium sulfate, and rotary evaporated. The reaction residue had δ^{119}Sn values in the NMR spectrum in CDCl$_3$ solution of -145.2, -142.0, -142.4, -139.8 (major), -130.2, -107.5, -101.6, -92.6 and –78.7ppm. The residue was separated, on a chromatotron, using hexane/ethyl acetate (2:1) as

eluent. The major fraction contained the triphenylstannylmethyl derivative of (*Z,Z*)-**3** (Ar = Ph, R = H). It was recrystallised from ethyl acetate; yield 23%, m.p. 170-174 °C. Anal. calcd. for $C_{62}H_{60}N_2O_8Sn$: C, 62.1; H, 5.1; N, 2.3. Found: C, 63.0; H, 5.5; N, 2.2%. IR (ν_{max}/cm^{-1}) 3297, 1676, 1507.

NMR Data are displayed in Tables 1 and 2.

X-Ray crystallography

Data were collected on an Enraf-Nonius CAD-4 diffractometer: data collection: CAD-4/PC.[27] Cell refinement: SET4 and CELDIM Software.[27] Data reduction: DATRD2 in NRCVAX94.[28] Program used to solve structure: SHELXS-97. [29] Program used to refine structure: SHELXL-97.[30] Preparation of material for publication: SHELXL-97 [30] and WordPerfect macro PRPKAPPA.[31] Diagrams were prepared with the aid of PLATON.[32]

The compound **4** was found to be unstable in the X-ray beam. The compound lies on an inversion centre in the crystal and has an overall *Z*-shape. H atoms were treated as riding atoms with C-H 0.93 to 0.98Å, N-H 0.86Å. Atoms C(7) to C(12) had anisotropic displacement parameter restraints applied to them. The C(14)-C(14)^a distance (symmetry operation: *a*: -x, 2-y, 1-z) was constrained to be 1.520(5)Å, this was the average distance found for this bond in the 3 molecules reported in **3** (Ar = Ph, R = H).[8] The largest peaks in the difference map were adjacent to the Hg atom at distances 1.11 to 1.06Å.

Crystal refinement data for the compound are listed in Table 5.

Table 5. Crystal refinement data for **4**

Molecular formula	$C_{32}H_{42}Cl_2Hg_2N_2O_{10}$
Formula weight	1086.76
T (K)	293(2)
Wavelength (Å) Mo *Ka*	0.71070
Crystal system	Triclinic
Space group	P-1
Cell dimensions	
a (Å)	8.088(3)
b (Å)	10.897(5)
c (Å)	11.486(4)
α (°)	66.75(3)
β (°)	88.85(3)
γ (°)	75.31(3)
Volume (Å^3)	896.1(7)
Z	1
Calculated density (Mg/m^3)	2.0139(16)
Absorption coefficient, mm^{-1}	8.762
F(000)	522
Crystal color	colourless
Crystal size (mm)	0.417x0.125x0.042
Theta range for data collection (°)	2.1 to 25.0
Index range	$-9 \le h \le 9$
	$0 \le k \le 12$
	$-12 \le l \le 13$
Refinement method	Full matrix least-squares on F^2
No. of reflections: no of parameters	3159, 219
Goodness-of-fit on F^2	0.99
Final R indices [I>2sigma(I)]	R = 0.094
	wR = 0.2458
Largest diff. peak and hole e.A^{-3}	2.80, -4.15

Supplementary material

Supplementary X-ray data for **4** are available from the Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK (fax: +44-1223-336033) on request, quoting the deposition number CCDC 153008.

Acknowledgement

The authors thank Professor George Ferguson, University of Guelph, Canada for the X-ray data collection. JLW thanks CPNq for support as a visiting professor at UFRJ.

References

1. Pierce, J.S.; Lunsford, C.D.; Raiford, R.W.; Rush, J.L.; Riley, D.W.; *J. Am. Chem. Soc.* **1951**, *73*, 2595.
2. Frankel, M.B. *US Patent 3.228.929*, **1966** [*CA*: *64*, *P9594*].
3. EniChem Anic S.P.A.: *Jpn. Kokai Tokkyo. J. Patent 62.249.958 (87.249.958)*, **1988** [*CA*: *109*, P92287].
4. Malhotra, S.L.; Naik, K.N.; MacKinnon, D.N.; Mayo, J.D.; Gagnon, Y.; Goodgrand, H.B.; *US Patent 57709737*, **1998** [*CA*: *128*, *P129292*].
5. Wheeler, E.L.; Jancis, E.H.; Gencarelli, R.A.; Barrows, F.H.; *US Patent 4.304.714*, **1982** [*CA*: *96*, *P86482*].
6. Dubois, G.E.; Zhi, B.; Roy, G.M.; Stevens, S.Y.; Yalpani, M.; *J. Chem. Soc., Chem. Commun.* **1992**, 1604.
7. Rammo, J; Schneider, H.-J.; *Liebigs Ann.* **1996**, 1957.
8. Ross, J.-N.; Wardell, J.L.; Low, J.N.; Ferguson, G.; *Acta Crystallogr.* **1996**, *C52*, 228.
9. Mohamadi, F.; Richards, N.G.J.; Guida, W.C.; Liskamp, R.; Lipton, M.; Caulfield, C.; Chang, G.; Hendrickson, T.; Still, W.C.; *J. Comput. Chem.* **1990**, *11*, 440.
10. Still, W.C.; Tempczyk, A.; Hawley, R.C.; Hendrickson, T.; *J. Am. Chem. Soc.* **1990**, *12*, 6127.
11. Eliel, E.L.; Wilen, S.H.; *Organic Stereochemistry*; Wiley: New York, 1994.
12. Wardell, J.L. In *Inorganic Reactions and Mechanisms*; Zuckerman, J.J. ed.; VCH Publications: New York, 1988, Vol. 11, 277.

13. Cremer, D.; Pople, J.A.; *J. Amer Chem. Soc*. **1975**, *97*, 1354.
14. Casas, J.S.; García-Tasende, M.S.; Sordo, J.; *Coord. Chem. Revs*. **1999**, *193-195*, 283.
15. Holloway, C.E.; Melnik, M.; *J. Organomet. Chem*. **1995**, *495*, 1.
16. Davies, A.G.; Wardell, J.L. In *Comprehensive Organometallic Chemistry*; Abel, E.W.; Stone, F.G.A.; Wilkinson, G. eds.; 2nd ed.; Pergamon Press: Oxford, 1995, Volume 3, p 135; Wardell, J.L. In *Comprehensive Organometallic Chemistry*, Wilkinson, G.; Stone, F.G.A.; Abel, E.W. eds.; 1st ed.; Pergamon Press: Oxford, 1982, Volume 1, p 863.
17. Grdenic, D.; *Quart. Rev*. **1955**, *19*, 303.
18. Canty, A.J.; Deacon, G.B.; *Inorg. Chim. Acta* **1980**, *45*, L225.
19. Batsanov, S.S.; *Russ. J. Inorg. Chem*. **1991**, *36*, 1694.
20. Huheey, J.E.; Keiter, E.A.; Keiter, R.L.; *Inorganic Chemistry*, 4th ed., Harper Collins: New York, 1993.
21. Zamora, F.; Sabat, M.; Janik, M.; Siethoff, C.; Lippert, B.; *Chem. Commun*. **1997**, 485.
22. Wardell, J.L.; Wigzell, J.McM.; *J. Organomet. Chem*. **1982**, *244*, 225.
23. Cox, P.J.; Doidge-Harrison, S.M.S.V.; Howie, R.A.; Wardell, J.L.; *J. Crystallogr. Spectrosc. Res*. **1991**, *21*, 735.
24. Ross, J.-N.; Wardell, J.L.; Low, J.N.; Ferguson, G.; *Acta Crystallogr*. **1996**, *C52*, 1961.
25. Allen, F. H., Kennard, O.; *Chem. Des. Autom. News* **1993**, *8*, 31.
26. see for example, Lo, K.M.; Ng, S.W.; Das, V.G.K.; *J. Cryst. Chem. Res*. **1997**, *12*, 933; Fu, F.Z.; Li, H.Y.; Zhu, D.S.; Fang, Q.X., Pan, H.A.; Tiekink, E.R.T.; Kayser, F.; Biesemans, M.; Verbruggen, I.; Willem, R.; Gielen, M.; *Organometallics* **1995**, *490*, 163; Jurkschat, K.; Schilling, J.; Mugge, C.; Tzschach, A.; Meunier-Piret, J.; van Meerssche, M.; Gielen, M.; Willem, R.; *Organometallics* **1988**, *7*, 38.
27. Enraf-Nonius; *CAD4-PC Software; Version 1.1*; Delft, Holland, 1992.
28. Gabe, E.J.; Le Page, Y.; Charland, J.-P.; White, P.S.; *J. Appl. Cryst*. **1989**, *22*, 384.
29. Sheldrick, G.M; *SHELXS-97; Program for the solution of crystal structures*; University of Göttingen, Germany, 1997.
30. Sheldrick, G.M.; *SHELXL-97; Program for the refinement of crystal structures*; University of Göttingen, Germany, 1997.
31. Ferguson, G.; *PRPKAPPA - a WordPerfect-5.1 macro to formulate and polish CIF format files from the SHELXL-97 refinement of Kappa-CCD data*; University of Guelph, Canada, 1999.
32. Spek, A.L.; *PLATON; Molecular Geometry and Graphics Program, January March 2000 version*; University of Utrecht, Holland, 2000.

Received: May 19, 2001
Published on the web: February 22, 2002