



Room 14-0551
77 Massachusetts Avenue
Cambridge, MA 02139
Ph: 617.253.5668 Fax: 617.253.1690
Email: docs@mit.edu
<http://libraries.mit.edu/docs>

DISCLAIMER OF QUALITY

Due to the condition of the original material, there are unavoidable flaws in this reproduction. We have made every effort possible to provide you with the best copy available. If you are dissatisfied with this product and find it unusable, please contact Document Services as soon as possible.

Thank you.

Some pages in the original document contain pictures, graphics, or text that is illegible.

Molecular Recognition of Adenosine Derivatives

by

Michael Morgan Conn
B.Sc., Chemistry and Biochemistry
University of Toronto

submitted to the Department of Chemistry in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

at the
Massachusetts Institute of Technology
February 1994

© 1994 Massachusetts Institute of Technology
All rights reserved

Signature of Author.....

Department of Chemistry
January 14, 1994

Certified by.....

Professor Julius Rebek, Jr.
Thesis Supervisor

Accepted by.....

Professor Glenn A. Berchtold

Chairman, Departmental Committee on Graduate Students

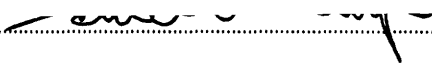
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

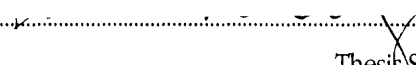
MAR 21 1994

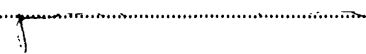
LIBRARIES

Science

This doctoral thesis has been examined by a Committee of the Department of Chemistry as follows:

Professor D. S. Kemp 
Chairman

Professor J. Rebek, Jr. 
Thesis Supervisor

Professor J. A. Stubbe 

MOLECULAR RECOGNITION OF ADENOSINE DERIVATIVES

by

Michael Morgan Conn

Submitted to the Department of Chemistry on January 14, 1994 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Chemistry

ABSTRACT

A multipurpose receptor module for 9-substituted adenines has been developed. The receptor makes use of imides derived from Kemp's triacid attached to 3,6-diaminocarbazole to provide simultaneous Watson-Crick and Hoogsteen base-pairing to an adenine stacked on the carbazole moiety. The combination of hydrogen-bonding and aromatic stacking interactions provides strong binding in the full range of solvents from non-polar organic to aqueous. The substituent on the secondary amine of carbazole (N⁹) can be varied to customize the function of the receptor:

A flexible tail whose conformation is induced by an additional hydrogen-bond to adenine N³ and aromatic stacking to the free face of the adenine. With an affinity for adenine of over -8 kcal/mol in chloroform, this receptor has the highest affinity for 9-ethyladenine yet reported.

A rigid phenyl substituent with a guanidinium group suspended from it for recognition of cyclic adenosine monophosphates in methanol. Molecular modeling studies led to the prediction that 2',3'-cyclic adenosine monophosphate would be bound preferentially on the basis of a more favorable salt-bridge geometry. This manifests itself as 0.5 kcal/mol stronger binding to 2',3'-cAMP than to 3',5'-cAMP.

An isophthaloyl substituent that can be used to construct a receptor for the adenosine trinucleotide by providing divergent attachment points for adenosine monophosphate receptors. Preliminary results on the transport of ApApA through liquid membranes suggests that the affinity of this receptor for the trinucleotide in organic solvents is so high that transport efficiency is compromised.

A biphenyl substituent for a self-replicating system designed to reduce the involvement of competing non-autocatalytic pathways. The high binding affinity of the adenosine module is designed to allow autocatalysis to occur in polar solvents, but complicates analysis of the system.

Molecular modeling studies and stochastic dynamics simulations were an integral part of the design and analysis process. Solution structures were deduced from control studies and both variable temperature and two-dimensional ¹H NMR experiments.

This receptor architecture was exploited for the determination of the strength of an intramolecular base-pair. From this, the strength of a hydrogen-bond in chloroform was deduced to be 1.5 kcal/mol.

Re-investigation of the complexation behavior of an earlier diimide receptor, one based on 2,7-diaminonaphthalene, led to discovery that 2:1 binding of adenine competes with 1:1 chelation. This behavior is *not* observed with diaminocarbazole-based receptors due to a lower propensity for aromatic stacking.

The syntheses were reported for both a structure-less receptor designed to be driven into its binding conformation by hydrogen-bonding and aromatic stacking interactions and a water-soluble autocatalytic system. Limited solubility prevented the analysis of these systems.

Thesis Supervisor: Dr. Julius Rebek, Jr.

Title: Camille Dreyfus Professor of Chemistry

ACKNOWLEDGEMENTS

First, the partners-in-crime. The chemist who initiated the 'scorpion' project was Ghislain Deslongchamps. His spirit and intellectual leadership have left an indelible impression on this group. The carbazole-diimide architecture was based on discussions with Javier de Mendoza, who has been an eager collaborator on the nucleotide transport studies. Neil Branda brought Ullman coupling into the group. His experimental temerity has opened up whole new avenues for us. Brian Murray managed to take our understanding of association constants and their determination to a new level, in the short time he was here. Yoko Kato's diligence and precision have been both refreshing and productive in our collaboration on molecular recognition in aqueous solution. And then, there's Ed. 'Sparky' has suffered objugation above and beyond the call of duty as my baymate and collaborator on the autocatalysis project. I am grateful for his expertise and creativity and for exercising impartiality and patience when I haven't.

Cecilia Andreu was responsible for all the nucleotide transport studies. Ivan Huc performed the vapor phase osmometry measurements on the self-complementary scorpion molecules. His ingenuous suggestion to use THF for nitro-carbazole hydrogenations saved me from my own cleverness and permitted synthesis of the trinucleotide receptor. Belinda Tsao developed the equations used to solve the titration involving simultaneous 2:1 binding and 1:1 chelation.

While any endeavor bears the mark of many influences, I would especially like to thank James Nowick, Ghislain Deslongchamps, and Pablo Ballester for their professional and personal tutelage. Their example has served to make me the chemist that I am today.

Learning is a long and convoluted process, but I would like to acknowledge Professors John Bunting, Bert Holland, and Fred Greene for their guidance through the years. I have been inspired by the spark I see in their eyes when they talk Chemistry.

Though they haven't the foggiest idea of what I do, I would like to thank my parents, Genevieve Conn and Charles Conn, as well as Babi and Dziadzi, for their support and encouragement and for giving me their drive for excellence, their work ethic and, of course, their genes. I would like to thank my brother, Grad, for a piece of advice that I have followed so far for almost fifteen years that "...you can *always* learn something." My kid sister, Meredyth, continues to help me appreciate the things that I'm not.

Finally, Julius. I am grateful for having the freedom to learn, pursue dumb ideas, blow money on the NMR, and become my own chemist. I've come away the better from our association. I'll always admire the creativity and fun you have with chemistry, but will never understand how you manage to kill so many floppy disks. So long and thanks for all the fish.

I would like to thank the National Science Foundation and the Natural Science and Engineering Research Council (Canada) for support in the form of predoctoral fellowships.

TABLE OF CONTENTS

ABSTRACT	5
ACKNOWLEDGEMENTS.....	7
TABLE OF CONTENTS	9
SYNTHETIC SCHEMES	13
LIST OF FIGURES	15
LIST OF TABLES.....	19
A. ADENINE BINDING: COMPLEXATION THROUGH INDUCED BINDING GEOMETRY.....	21
1. Introduction.....	21
2. Diimide Clefs.....	22
2.1 Extraction Studies.....	25
2.2 Induced Binding Geometry	26
2.3 Binding Studies.....	30
3. Determining Hydrogen-Bond Strength	37
4 Reinvestigation of Naphthyl-Diimide Receptors	40
5. Monoimide Receptors	48
6. Experimental Section.....	50
6.1 General.....	50
6.2 Synthesis.....	51
6.3 Extractions	61
6.4 Titrations.....	61
6.5 NMR Spectroscopy.....	62
6.6 Molecular Modeling.....	63
B. NUCLEOTIDE RECOGNITION.....	65
1. Introduction.....	65
2. Selective Cyclic Monophosphate Binding	67
3. Trinucleotide Recognition	71
4. Nucleotide Transport	74
4.1 Mononucleotides	75
4.2 Dinucleotides	77
4.3 Trinucleotides	78
5. Experimental	79

5.1 Synthesis.....	79
5.2 Titrations.....	84
5.3 Modeling.....	85
5.4 Transport.....	85
C. SELF-REPLICATING MOLECULES: A SECOND GENERATION.....	87
1. Background.....	87
2. Self-Complementary Molecules.....	92
3. Autocatalysis.....	99
4. Water-Solubilization of an Autocatalytic System.....	113
5. Experimental.....	115
5.1 Synthesis.....	115
5.2 Spectroscopy.....	124
5.3 Titrations.....	124
5.4 Molecular Modeling.....	125
5.5 Vapor Phase Osmometry.....	125
5.6 Kinetics.....	125
D. BIBLIOGRAPHY.....	127
E. SPECTRA.....	133

SYNTHETIC SCHEMES

Scheme A.2.1. Synthesis of diimide cleft.

Scheme A.2.2. Synthesis of cap fragment.

Scheme A.2.3. Synthesis of tail fragments.

Scheme A.3.1. Synthesis of carbazole monoimide.

Scheme A.3.2. Synthesis of water-soluble carbazole monoimide.

Scheme A.5.1. Synthesis of induced-binding cleft.

Scheme A.5.2. Synthesis of water-soluble induced-binding cleft.

Scheme B.2.1. Synthesis of phenylscorpion.

Scheme B.2.2. Assembly of guan-phenylscorp.

Scheme B.3.1. Synthesis of isophthaloylscorpion.

Scheme C.2.1. Synthesis of methylscorpion template.

Scheme C.2.2. Synthesis of tris(hydroxymethyl) methylscorpion template.

Scheme C.2.3. Synthesis of phenylscorpion template.

Scheme C.3.1. Synthesis of N-(4-(4'-(methoxycarbonyl)phenyl)phenyl)-3,6-dinitrocarbazole.

Scheme C.3.2. Synthesis of tripropyl biphenylscorpion template.

Scheme C.3.3. Synthesis of control for the biphenylscorpion autocatalytic system.

Scheme C.4.1. Synthesis of tris(hydroxymethyl) biphenylscorpion template

Scheme C.4.2. Masked deprotection of tris((benzyloxy)methyl) biphenylscorpion.

LIST OF FIGURES

- Figure A.1.1. The hydrogen-bonding modes of adenine.
- Figure A.2.1. Schematic of scorpion cleft geometry.
- Figure A.2.2. Synthesis of scorpion clefts.
- Figure A.2.3. Additional diimide clefts.
- Figure A.2.4. Geometry and proton identification of the complex of **A24** and 9-ethyladenine.
- Figure A.2.5. Variation of NOE with tumbling rate.
- Figure A.2.6. Lowest energy conformation of **A24** complexed with 9-ethyladenine (stereoview).
- Figure A.2.7. Possible binding modes for guanosine derivatives.
- Figure A.4.1. The different binding modes of naphthalene-diimide receptor.
- Figure A.4.2. Naphthyl-based imide clefts.
- Figure A.4.3. Titration of 9-ethyladenine with naphthyl diimide does not fit 1:1 binding model.
- Figure A.4.4. Mixed 2:1, 1:1 chelation.
- Figure A.4.5. Titration of 9-ethyladenine with diimide **A36** fits combined 1:1/2:1 model.
- Figure A.4.6. Superimposed **A5** and **A36** complexes of 9-ethyladenine.
- Figure A.4.7. Adenine superimposed on carbazole (left) and naphthalene (right) as in the complexes.
- Figure A.5.1. Schematic of symmetric bis(imide)-adenine binding.
- Figure A.5.2. Predicted geometry of mono-imide clefts induced by binding (stereoview).
- Figure B.1.1. First-generation nucleotide transport agents.
- Figure B.1.2. Adenosine nucleotide targets.
- Figure B.2.1. Predicted structure of **B9** complexed with 2',3'-cAMP (stereoview).
- Figure B.2.2. Predicted structure of **B9** complexed with 3',5'-cAMP (stereoview).
- Figure B.3.1. An adenosine trinucleotide receptor.
- Figure B.3.2. Assembly of trinucleotide receptor **III**.
- Figure B.4.1. U-tube apparatus for liquid membrane transport studies.
- Figure B.4.2. Liquid membrane transport concentration profile.
- Figure B.4.3. Transport control.

-
- Figure B.4.4. Dinucleotide carrier **V**.
- Figure B.4.5. Controls for trinucleotide transport.
- Figure C.1.1. Schematic of self-complementary template-based autocatalysis.
- Figure C.1.2. Self-replication of DNA trimers.
- Figure C.1.3. The first abiotic self-replicating system.
- Figure C.1.4. Controls for catalysis by self-replication.
- Figure C.1.5. Schematic of lactam catalysis of tetrahedral intermediate breakdown.
- Figure C.1.6. Control for fit-constraints in self-replication.
- Figure C.1.7. Intracomplex reaction in the naphthoyl self-replicating system.
- Figure C.1.8. Intracomplex reaction in the biphenyl self-replicating system.
- Figure C.2.1. Predicted geometry of the intramolecular conformation of **C1** (stereoview).
- Figure C.2.2. Simulation of imide chemical shift with concentration.
- Figure C.2.3. Predicted structure of folded **C5** (stereoview).
- Figure C.2.4. A molecular jackknife.
- Figure C.2.5. Predicted structure of the molecular jackknife (stereoview).
- Figure C.3.1. The complex between the biphenylscorpion and aminoadenosine.
- Figure C.3.2. The autocatalytic reaction.
- Figure C.3.3. Catalytic enhancement of autocatalysis in chloroform/THF mixtures.
- Figure C.3.4. Predicted structure of the autocatalytic tetrahedral intermediate (stereoview).
- Figure C.3.5. Predicted structure of a tetrahedral intermediate in monoimide autocatalysis (stereoview).
- Figure C.3.6. Reaction rate of autocatalysis in 13% THF/chloroform.
- Figure C.3.7. The intrinsic background reaction for autocatalysis.
- Figure C.3.8. Control studies with the pentafluorophenyl ester of **C17**.
- Figure C.3.9. Additional control for autocatalysis.
- Figure C.3.10. The structural elements of the template are not catalysts, individually.
- Figure C.3.11. Variation of free substrate and termolecular complex concentration with starting concentration.

LIST OF TABLES

Table A.2.1. Equivalents of nucleoside derivatives extracted from aqueous solution

Table A.2.2. Intermolecular NOESY correlations in CD₂Cl₂ solution.

Table A.2.3. Estimates of association constants of receptors with 9-ethyladenine in CDCl₃.

Table A.2.4. Estimates of association constants of receptors with 9-ethyladenine in CD₃OD.

Table A.2.5. Selectivity of **A24** for various nucleoside derivatives in CD₃OD.

Table B.2.1. Association constants of receptor **B9** with cyclic AMP (Na⁺ salt) in CD₃OD.

Table B.4.1. Transport of mononucleotides by carrier **I**.

Table B.4.2. Transport of mononucleotides by **B10**.

Table B.4.3. Transport of dinucleotides by carrier **V**.

Table B.4.4. Transport of ApApA (10 mM).

Table C.3.1. Reaction rates of aminoadenosine with pentafluorophenyl ester of **C14** in 13% THF/CHCl₃.

Table C.3.2. Reaction rates of aminoadenosine with pentafluorophenyl ester of **C17** in 13% THF/CHCl₃.

Table C.3.3. Control reactions of aminoadenosine with pentafluorophenyl ester of **C14** in 13% THF/CHCl₃.

A. ADENINE BINDING: COMPLEXATION THROUGH INDUCED BINDING GEOMETRY

1. INTRODUCTION

Molecular recognition of nucleic acid components has attracted considerable attention in recent years, and all of the naturally-occurring nucleosides have been the focus of such studies. Successful strategies typically employ receptor surfaces that simultaneously interact with the target through hydrogen bonding and aromatic stacking, while such additional features as salt bridges assist in the recognition of phosphate derivatives.¹⁻¹⁰ In aqueous solution, interactions among the nucleosides are weak and primarily involve aggregation through base-stacking forces¹¹ whereas hydrogen bonding dominates in nonpolar organic solvents.^{12,13} Correspondingly, molecular recognition studies have predominately been carried out in organic solvents in order to exploit hydrogen-bonding as the major contributor to binding.

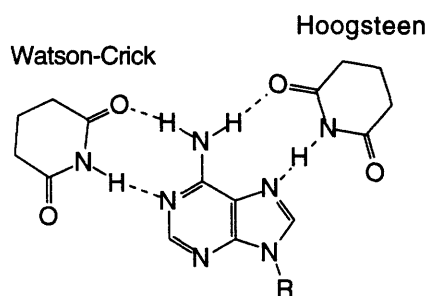


Figure A.1.1. The hydrogen-bonding modes of adenine.

Because of the vital role it plays as a component of a number of important biological cofactors, adenosine has been the focus of especial attention in molecular recognition studies. In addition to the aromatic system, adenosine presents two separate hydrogen-bonding surfaces for recognition. Base pairing with these two faces results in two modes of binding – the Watson-Crick mode typically seen in DNA and the Hoogsteen mode. A number of widely different receptor designs are possible with these multiple binding modes, including macrocycles,² tweezers,⁴ and clefts.¹⁴ Refinement of these designs has resulted in receptors with very high affinity for adenosine derivatives. This laboratory has previously reported¹⁵ the synthesis of a cleft based on 2,7-diaminonaphthalene able to complex in both the Hoogsteen and Watson-Crick modes at once while simultaneously providing additional stabilization through aromatic stacking. This receptor binds to 9-ethyladenine with an affinity of 10^5 M^{-1} (-6.7 kcal) in CDCl_3 .¹⁵ Preliminary reports with extraction¹⁶ suggest that this is the most efficient receptor to date.

The significant advantage of biological receptors over synthetic ones is their ability to strongly bind substrates in aqueous solution. This is due to hydrophobic binding pockets within the structure of the protein. The entropy gained from the release of water from these pockets is an important driving force for complexation.¹⁷ A practical strategy for the construction of binding pocket models would involve modification of the skeleton of

the most efficient pre-existing receptor. Ideally, this would be synthetically simple and general to allow rapid assembly of clefts for structure-function studies. While 2,7-diaminonaphthalene does not lend itself to this easily, a receptor based on 3,6-diaminocarbazole allows facile functionalization of the aromatic moiety while maintaining optimum orientation of the hydrogen-bonding surfaces. We have recently reported the use of this binding motif, coupled with a guanidinium unit, for the complexation of the adenine dinucleotide¹⁸ and cyclic adenosine monophosphates.¹⁹

2. DIIMIDE CLEFTS

The bulk of this section has previously appeared in print.²⁰

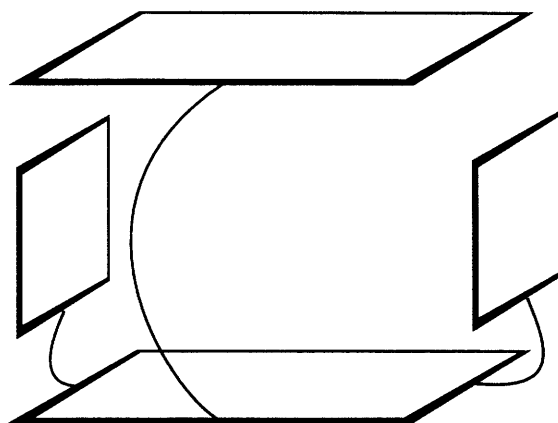
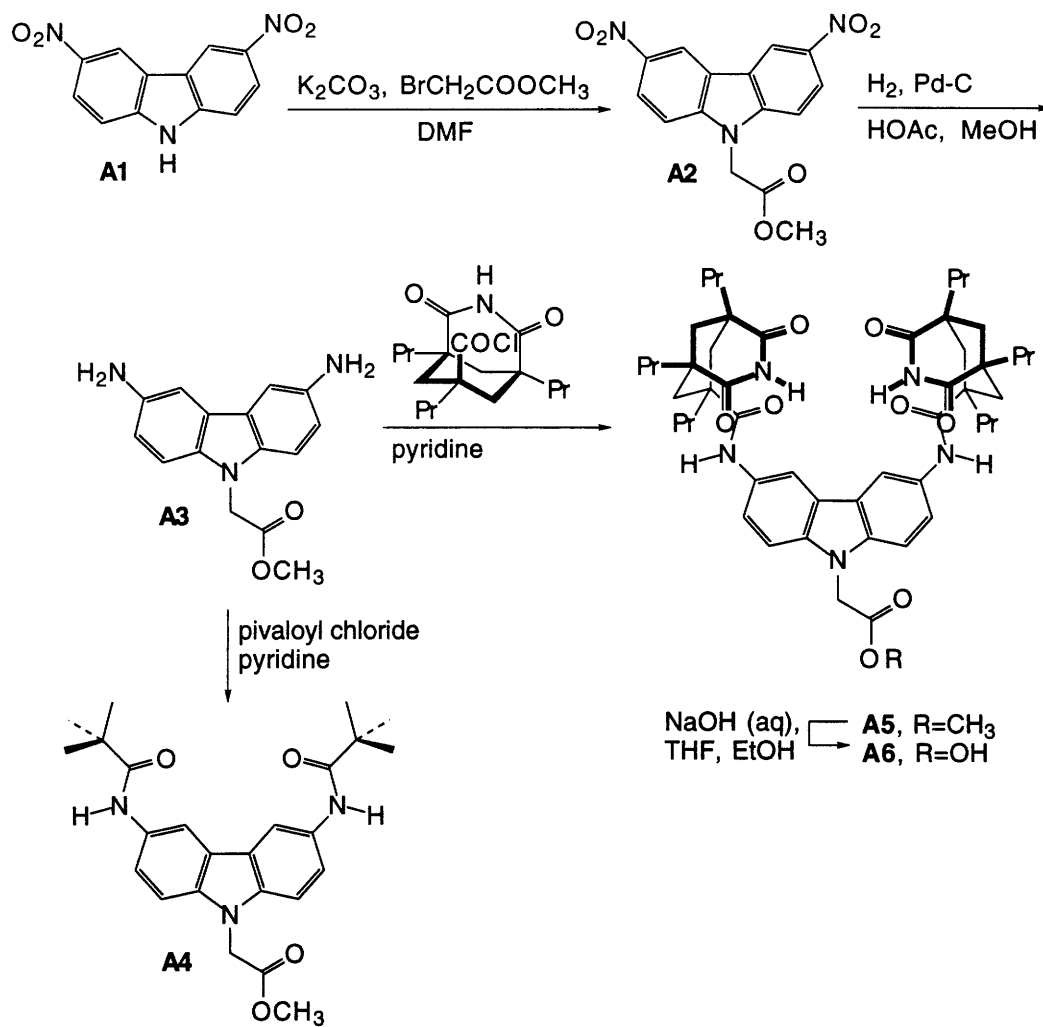


Figure A.2.1. Schematic of scorpion cleft geometry.

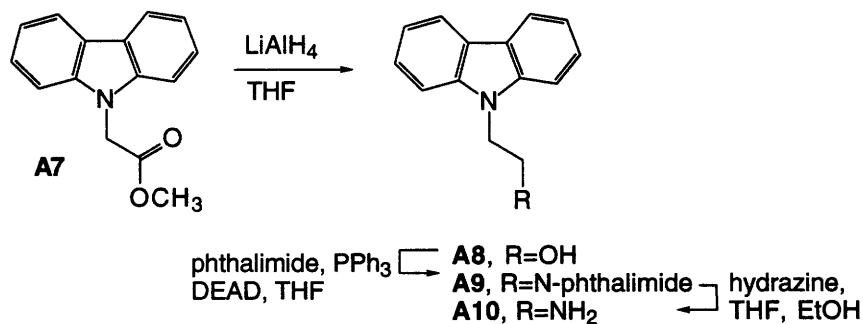
In this work, we introduce a family of receptors designed to bind adenosine derivatives within a pocket formed through induced fit. The architecture of these receptors consists of a 3,6-diaminocarbazole-based receptor functionalized with a pendant tail which can stack on top of an adenosine within the hydrogen-bonding cleft. This 'scorpion-like'²¹ binding geometry is shown schematically in Figure A.2.1 and provides a mechanism of sequestering the bound nucleoside from the bulk solvent. This work discusses the behavior of these receptors in polar and nonpolar organic solvents. In aqueous solution, such an architecture could both protect the hydrogen bonding surface from competition with water and introduce a favorable entropic contribution to binding through the release of bulk water.

The receptors in this study were readily assembled using standard synthetic methods. Scheme A.2.1 summarizes the assembly of the core adenine-binding unit. Dinitrocarbazole (**A1**) was alkylated and reduced to give diaminocarbazole **A3**. This hydrogenation step was extremely difficult and proceeded with erratic yield. The diaminocarbazole is a highly air- and light-sensitive compound and was used immediately following reduction. The hydrogenation conditions were designed to stabilize the diamine as it formed to prevent decomposition, by protonation with acetic acid. However, in later studies (Sections A.3, A.5, B and C) hydrogenation in pure THF, that is in the *absence* of alcoholic solvent, resulted in significantly greater stability of

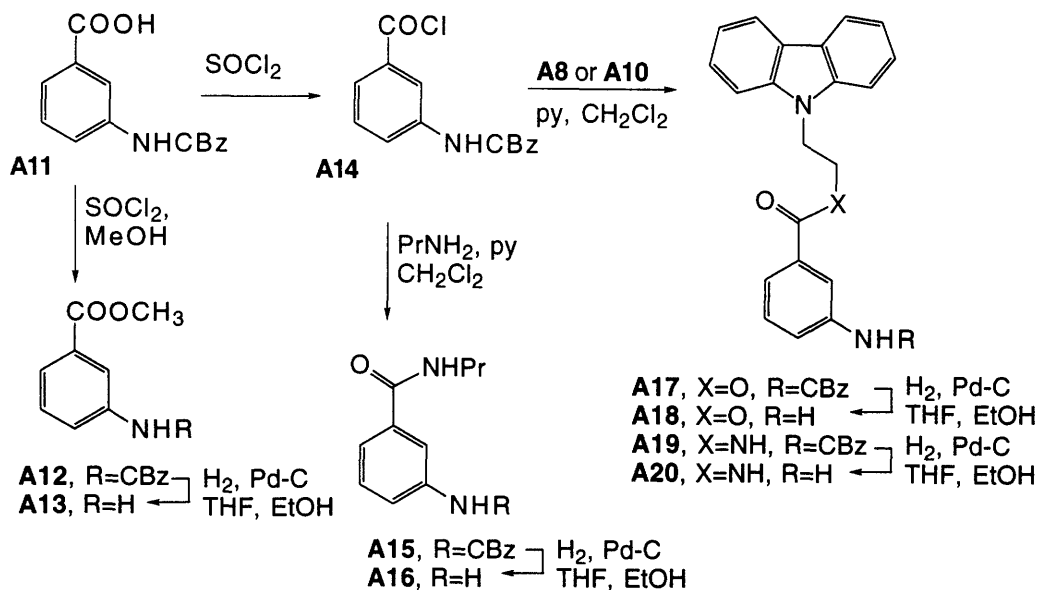
the diamine. Acylation of the diamine with the tripropyl derivative of the Kemp's triacid imide acyl chloride,²² gave the diimide cleft **A5** after considerable purification. Hydrolysis yielded the adenine binding unit **A6** ready for condensation with the tail subunits.



Scheme A.2.1. Synthesis of diimide cleft.



Scheme A.2.2. Synthesis of cap fragment.



Scheme A.2.3. Synthesis of tail fragments.

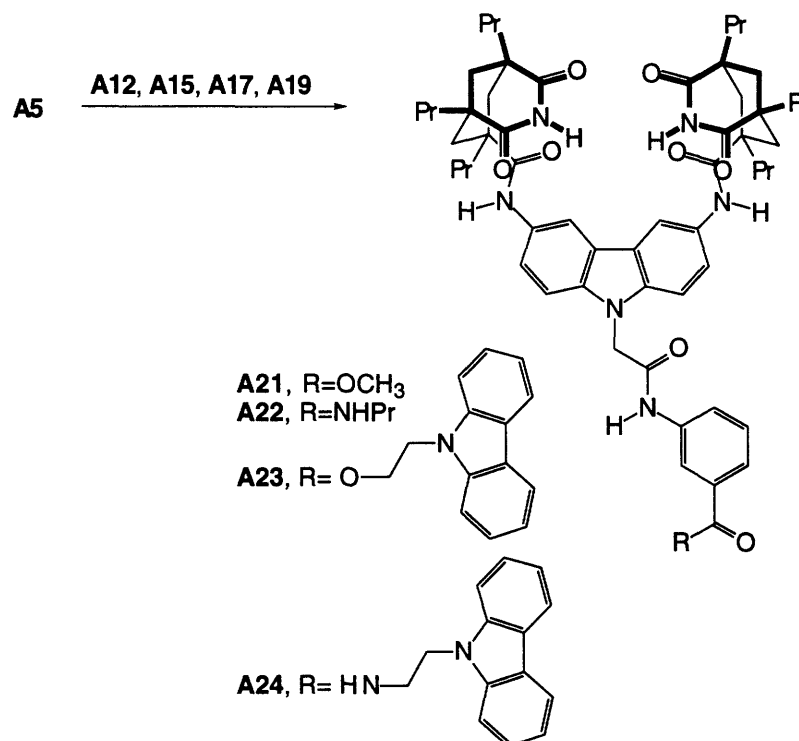


Figure A.2.2. Synthesis of scorpion clefts.

The cap fragments were synthesized by alkylation of carbazole, followed by reduction to alcohol **A8** (Scheme A.2.2). Mitsunobu inversion²³ of the alcohol with phthalimide, followed by hydrazinolysis gave amine **A10**.

The tail of the receptor is based on 3-aminobenzoic acid. The synthetic manipulations required for the assembly

of the tail are summarized in Scheme A.2.3. Condensation of the tail fragments with diimide **A5** gives the fully articulated receptors **A21-A24** (Figure A.2.2). Receptors **A21-A23** are incremental derivatives of **A24** that feature one-by-one removal of various elements expected to contribute to binding. These are analogous to site-specific protein mutations designed to elucidate the contribution of individual elements to binding affinity.^{24,25}

A set of additional clefts was assembled and their association constants measured by Ghislain Deslongchamps to evaluate additional contributions to binding (Figure A.2.3). The tail of receptor **A27** is longer than that of **A24** by one carbon and features inverted directionality of the distal amine group.

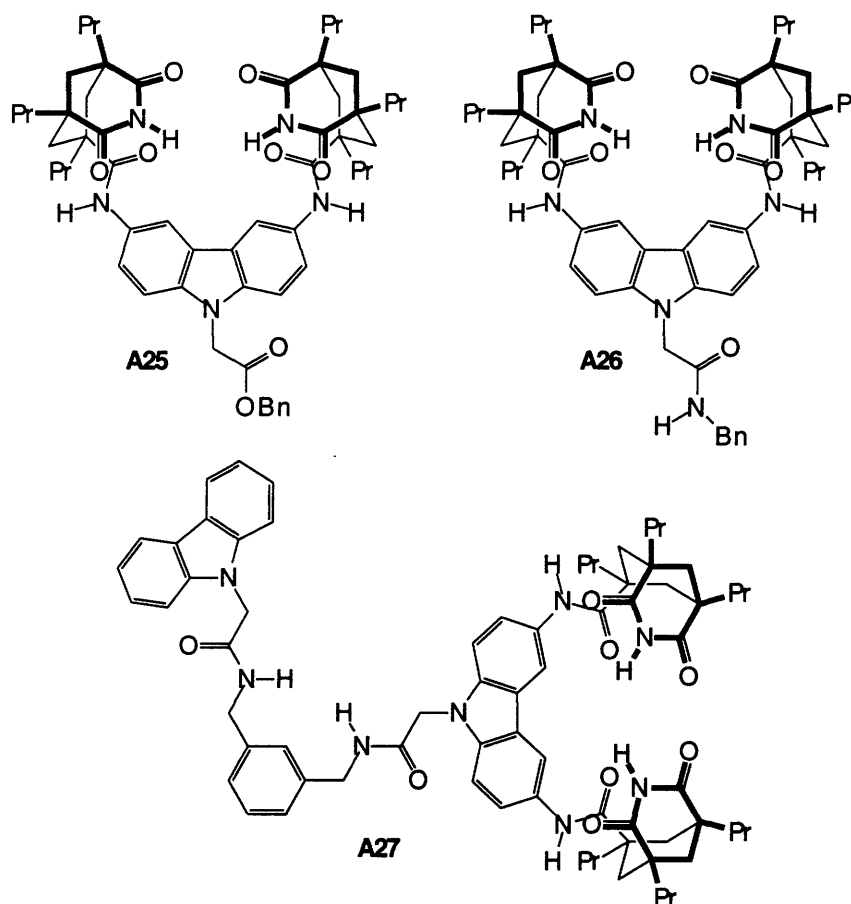


Figure A.2.3. Additional diimide clefts.

2.1 Extraction Studies

Adenosine-binding clefts based on Kemp's triacid imide have been previously used for transport of nucleoside derivatives across liquid membranes.¹⁶ The receptors in this study were examined to evaluate the affinity of these receptors for adenosine derivatives at the chloroform-water interface. Receptors **A24** and **A27** are able to extract up to a *full equivalent* of adenosine derivatives from aqueous solution into chloroform (Table A.2.1).

Extraction efficiency depends strongly on the water-chloroform partition coefficient of the target.²⁶ This dependence manifests itself as a decrease in extraction with increasing hydrophilicity (Table A.2.1) – a pattern

which mirrors that from our earlier study, despite the higher binding affinity of the receptors in this work. In this case, it seems that formation of a binding pocket around the purine base is not sufficient to overcome the unfavorable effect of partition coefficients on extraction.

nucleoside	A24	A27	none
adenine	1.12	1.00	0
2-deoxyadenosine	0.64	0.80	-0.1
adenosine	0.3	0.30	0
guanosine	-	0	-
3',5'-cAMP	-	0	-

Table A.2.1. Equivalents of nucleoside derivatives extracted from aqueous solution

2.2 Induced Binding Geometry

In chloroform, NMR spectra of the larger scorpions (**A21-A24**, **A27**) are extremely broadened and show concentration dependence. Low temperature NMR (-53 °C) in CDCl₃ failed to resolve any identifiable species. In contrast, spectra in CD₃OD are well-resolved at room temperature, indicating that the broadening seen is a result of the slow tumbling of intermolecular aggregates formed through hydrogen bonding.

Addition of 9-ethyladenine to a 10 mM solution of scorpion receptor **A24** in CDCl₃ causes a dramatic conformational reorganization, giving a single species by NMR (Figure E.1). The substrate-receptor complex exists as a pair of exchanging enantiomers, based upon the position of the N⁹-substituent of adenine relative to the pendant tail of the receptor (*i.e.* to the right versus the left of the tail – Figure A.2.4). Because of this, the chemically equivalent protons become diastereotopic (magnetically nonequivalent) in the complex. The racemization process requires at least partial opening of the complex, so the bound adenine can leave and re-enter in the opposite orientation. However, the protons on the carbazole lid can exchange by a rotation without adenine leaving the cleft. NMR signals due to these protons are noticeably sharp.

Racemization can be easily slowed at subambient temperature (Figure E.2). The barrier to conversion between diastereotopic protons at the coalescence temperature can be estimated from the maximal frequency difference in the signals, according to A.2.1 and A.2.2.²⁷ This frequency difference can be approximated as 110 Hz by the difference observed at low temperature (-10 °C). The rate of exchange at the coalescence temperature (approximately 10 °C), then, is 244 s⁻¹ ($\Delta G^\ddagger \sim 13.6$ kcal/mol). No unbound 9-ethyladenine is observed at this temperature.

$$k_r = \frac{\pi}{\sqrt{2}} \Delta\nu \quad (\text{A.2.1})$$

$$\Delta G^\ddagger = RT_c \ln\left(\frac{kT_c}{k_r h}\right) \quad (\text{A.2.2})$$

In addition to the splitting of the diastereotopic proton signals, the two imide protons display different chemical

shifts at low temperatures, reflecting the different chemical environments of the Watson-Crick and Hoogsteen modes. Another significant feature evident is the downfield movement of the amide proton NH^2 , suggesting its participation in hydrogen-bonding with N^3 of the 9-ethyladenine.

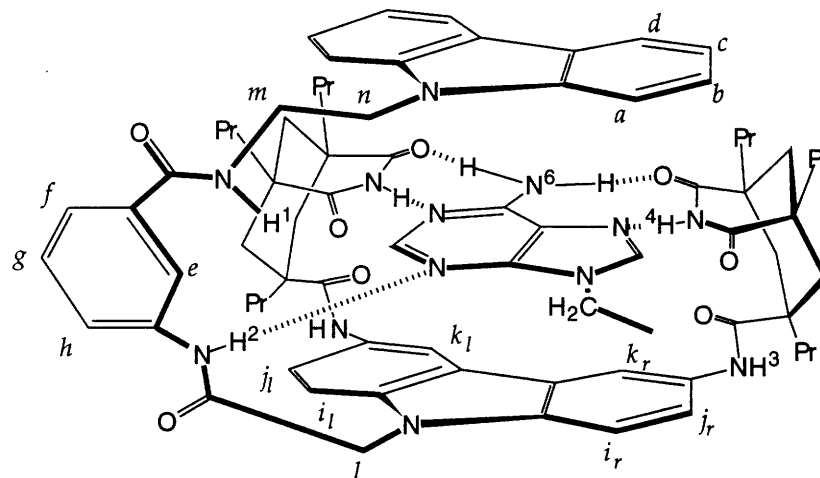


Figure A.2.4. Geometry and proton identification of the complex of **A24** and 9-ethyladenine.

Two-dimensional ^1H NMR studies in CD_2Cl_2 were used to obtain information on the geometry of the one-to-one complexes of **A24** and **A27** with 9-ethyladenine. Peak assignments were made using cross-correlations in TOCSY²⁸ (Figure E.3) and intramolecular correlations in NOESY²⁹ experiments (Figure E.4). The *intermolecular* contacts observed seen for the complexes of **A24** and **A27** are summarized in Table A.2.2. For both of these complexes, there is an NOE correlation between the imide protons and both adenine aromatic protons (C^8H and C^2H). This is direct evidence of the chelation of adenine within the cleft. In addition, there are through-space correlations between protons on the adenine and in the tail of the receptor. Since there is no pre-organization of the conformation of the receptor prior to binding, the tail can explore a wide range of conformational space. That the adenine CH_2 and protons on the tail are in close proximity to each other indicates that there is a fold where these protons are driven up close to the adenine nucleus. Figure A.2.4 is a tantalizing interpretation of this. We propose, in fact, that amide NH^2 is hydrogen-bonded to adenine N^3 based on these through-space correlations. This is supported by the large chemical shift change in the ^1H NMR spectrum of the signal due to this amide when the temperature is reduced.

A24 and 9-ethyladenine	A27 and 9-ethyladenine
NH^4 and C^8H	NH^4 and C^8H
NH^4 and C^2H	NH^4 and C^2H
H_e and CH_2	H_e and CH_2
NH^2 and CH_2	NH^2 and CH_2
H_h and CH_2	

Table A.2.2. Intermolecular NOESY correlations in CD_2Cl_2 solution.

No NOE correlations could be observed between any of the aromatic rings or, in particular, to the topmost carbazole. Rather than a failure of the complex to assume the geometry of Figure A.2.4, this is due to a limitation of the experiment imposed by the size and conformational dynamics of the complex. The nuclear Overhauser effect is a function of the frequency of NMR spectrometer, ω , and the correlation time, τ_c , of the molecule (Figure A.2.5).³⁰ At a given frequency, the observed NOE is a function of the tumbling rate of the molecule. For small, rapidly tumbling molecules, τ_c is very small and $\omega\tau_c$ is much less than one. As spectral lines are usually narrow from rapid averaging of signals, this is referred to as the extreme narrowing limit. The maximum possible NOE in this situation is +50%. For large molecules, $\omega\tau_c > 1$. This is referred to the extreme broadening limit and the maximum NOE is -100%. Molecules whose tumbling rate is near the crossover point ($\omega\tau_c = 1.12$) will show NOE enhancements that are sensitive to field strength, temperature, or solvent viscosity.

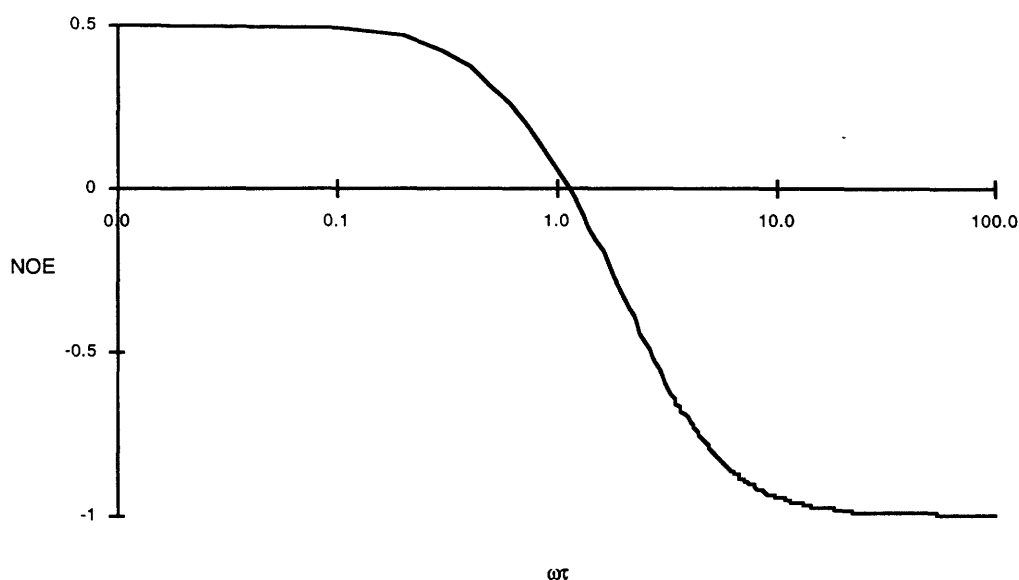


Figure A.2.5. Variation of NOE with tumbling rate.

In CD_2Cl_2 , all of the NOE correlations observed are positive, indicating that the tumbling rate of the complex is that of a small molecule. The size of the complex (1400 amu) and the breadth of the ^1H NMR signals suggest however that the complex is approaching the intermediate region between small and large molecule behavior. For instance, when the complex is dissolved in CDCl_3 , only *intramolecular* correlations can be observed – the less-favored intermolecular pathways were not observed. This strongly suggests that in this solvent the molecule has moved closer to the boundary between the extreme broadening and extreme narrowing limits. Rotating frame NOESY experiments (ROESY³¹) have been used to give stronger through-space correlations for molecules in this intermediate range, but could not be used. The ROESY experiment is complicated by several additional relaxation pathways including homonuclear Hartmann-Hahn correlations.^{30,32-34} In this highly aromatic system, it was impossible to separate ROE interactions from these interfering TOCSY-like correlations.

Despite the absence of interaromatic NOE's, indirect evidence for stacking of the cap over bound adenine was obtained from a titration of scorpion **A24** with 9-ethyladenine (a 'forward' titration) in CD₃OD. The proton NMR signals due to the carbazole cap showed significant (0.05-0.1 ppm) upfield shifts during the titration. Since control studies show that medium effects cause *much* smaller shifts (section A.2.3), this is consistent with complexation-induced stacking of the carbazole on bound adenine. This same study, by way of downfield shifts of carbazole proton *k*, also indicates inward orientation of the carbonyls of the amides that link the carbazole and Kemp's moieties.

The variable temperature studies of the complex of **A24** and 9-ethyladenine in CDCl₃ showed that considerable stiffening of the complex occurs at -50 °C. It should be possible to cause the molecule to shift well into the extreme broadening range at this temperature to enhance the observed NOE in NOESY experiments. The small driving force for aromatic stacking in non-polar organic solvents makes the stacked geometry only one of multiple stable conformations for the carbazole cap. By reducing the temperature, the conformational dynamics are reduced such that the residence time of the cap in this geometry is significantly increased.

At -50 °C, all of the observed NOE correlations within the complex of **A24** and 9-ethyladenine are large and negative -- as typically seen for large molecules at the extreme broadening limit. The complexation-induced diastereotopicity of chemically equivalent protons makes these spectra extremely complicated; only partial peak assignment of the ¹H NMR spectrum at this temperature (Figure E.2a) is possible from TOCSY (Figure E.5) and intramolecular NOESY correlations (Figure E.6). The conformational dynamics of the complex have been so severely limited at this temperature, in fact, that chemically equivalent protons on the bottom carbazole have different chemical shifts by virtue of being under either the imidazole ('right') or pyrimidine ('left') rings of the bound 9-ethyladenine.

A number of through-space correlations are observed that are consistent with the binding geometry predicted in Figure A.2.4. There are correlations between the imide protons and proton *k*, on the lower carbazole, as well as to proton *d*, on the *upper* carbazole, consistent with a 'three-tiered' geometry. Through-space correlations can be observed from CH₂ of 9-ethyladenine both to protons *i_r* and *j_r* but only very weakly to *i_l* or *j_l*. The CH₂ signal also shows cross-relaxation to proton *a* on the upper carbazole ring and to NH². The adenine aromatic protons show preferential through-space correlation to their corresponding superimposed carbazole protons. That is, C⁸H shows a strong NOE with *j_r* but only very weakly with *j_l*. C²H shows strong correlation to *j_l*, but no NOE with *j_r*. In addition, proton C²H shows an NOE to amide proton NH¹; there is also a weak CH₂ to NH¹ correlation. This indicates that NH¹ may also participate in a hydrogen-bonding interaction with adenine N³.

In addition, the two imide signals can be assigned as corresponding to the Watson-Crick and Hoogsteen hydrogen-binding modes. The imide at 13.4 ppm shows correlations to the adenine C²H proton, to proton *i_l* and a strong correlation to *j_l* and *k_l*. This is consistent with the Watson-Crick base-pairing mode. The imide at 13.65 ppm shows correlation to *j_r* and to the adenine C⁸H proton, consistent with the Hoogsteen-bound mode.

The complexity of the spectrum at this temperature does not entirely rule out the possible presence of multiple conformations of the upper carbazole section. The two signals at 6.7 ppm and 6.77 ppm both show crosspeaks consistent with that expected for amide NH¹, but only the proton at 6.7 ppm was identified as such.

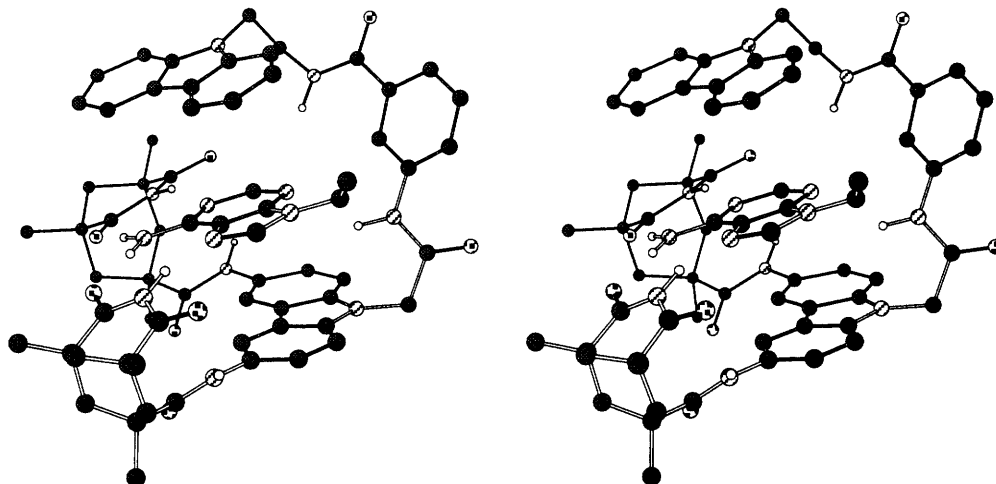


Figure A.2.6. Lowest energy conformation of **A24** complexed with 9-ethyladenine (stereoview).

These structural features were modeled through a stochastic dynamics simulation at 300 K of the complex of **A24** and 9-ethyladenine using chloroform solvation (see Experimental Section for details). Throughout the simulation, the adenine remains bound between both imides and shows hydrogen bonding to one or both of the amides in the tail. As the simulation proceeded, the carbazole cap moved away from the adenine (about 31 ps into the simulation), simulating the observed solution behavior. The lowest energy conformer, generated by minimization of structures periodically sampled from the simulation, is shown in Figure A.2.6. This structure is consistent with all experimental data -- chelated hydrogen bonding of adenine, hydrogen bonding to N³, and stacking of the carbazole lid. The flexibility of these receptors is such that it is unlikely that there is only one solution conformation of the complex. There are, in fact, a number of low energy conformations from the dynamics simulation. Rotation of a Kemp's imide moiety so that the adenine is bound in a bifurcated sense¹⁴ on one edge and tilting of the adenine so that hydrogen bonding can occur to NH¹ constitute the major forms of conformational variation. Conformations that show the carbazole lid *not* stacking onto adenine lie more than 20 kJ/mol above the lowest energy conformer.

2.3 Binding Studies

The choice of 3,6-diaminocarbazole as the scaffolding from which to build receptors involves some sacrifice in binding ability. There is an intramolecular hydrogen bond between the two imide groups which must be broken for binding to occur. Molecular modeling indicates that the curvature of the carbazole groups allows the two imides to interact with each other by rotating into the plane of the aromatic surface. Indeed, the imide protons of **A5** in CDCl₃ are shifted downfield (10.7 ppm) relative to unbound imide signals (7.6 ppm¹⁴) in the proton NMR spectrum. This behavior is not observed with the naphthalene diimide receptor as the geometry of

2,7-diaminonaphthalene does not allow close approach of the imide groups.

Binding constants were determined from ^1H NMR titrations, through the incremental addition of a solution of 'guest' to a solution of 'host' while following the chemical shift of protons on the host molecule. The 'host' and 'guest' are defined solely by their status in the titration protocol. In these cases, the 'host' is the adenosine derivative and the 'guest' is the receptor in question. This 'reverse' titration protocol was performed because of the extreme broadening of the ^1H NMR spectra of the larger receptors examined. The chemical shift of the host as a function of concentration of guest added forms a rectangular hyperbola known as the 1:1 binding isotherm.²⁶ From the equilibrium defined in A.2.1, the equilibrium constant can be defined as in A.2.2. The corresponding mass balance equations are shown in A.2.3 and A.2.4.



$$K_a = \frac{[\text{HG}]}{[\text{H}][\text{G}]} \quad (\text{A.2.2})$$

$$H_t = [\text{H}] + [\text{HG}] \quad (\text{A.2.3})$$

$$G_t = [\text{G}] + [\text{HG}] \quad (\text{A.2.4})$$

H is the 'host', G is the 'guest', and HG is the noncovalent complex between the two. H_t and G_t are the total concentrations of the two species in solution. The observed chemical shift of protons on the host is a weighted average of the chemical shifts of the free and bound host. This is mathematically expressed in Equation A.2.5.

$$\delta = \delta_H \frac{[\text{H}]}{H_t} + \delta_{\text{HG}} \frac{[\text{HG}]}{H_t} \quad (\text{A.2.5})$$

By substituting for the concentration of free host from Equation A.2.3, the chemical shift can be expressed solely in terms of the concentration of the complex, as in A.2.7.

$$\delta = \delta_H \left(1 - \frac{[\text{HG}]}{H_t} \right) + \delta_{\text{HG}} \frac{[\text{HG}]}{H_t} \quad (\text{A.2.6})$$

$$\delta = \delta_H + (\delta_{\text{HG}} - \delta_H) \frac{[\text{HG}]}{H_t} \quad (\text{A.2.7})$$

The concentration of complex can be expressed as a function of the total host and guest equations by substitution of the mass balance equations A.2.3 and A.2.4 into the equilibrium expression A.2.2 to give:

$$K_a = \frac{[\text{HG}]}{(H_t - [\text{HG}])(G_t - [\text{HG}])} \quad (\text{A.2.8})$$

Simplification of this gives a quadratic expression,

$$[\text{HG}]^2 - \left(H_t + G_t + \frac{1}{K_a} \right) [\text{HG}] + H_t G_t = 0 \quad (\text{A.2.9})$$

and the concentration of complex can be expressed according to Equation A.2.10.

$$[HG] = \frac{\left(H_t + G_t + \frac{1}{K_a}\right) + \sqrt{\left(H_t + G_t + \frac{1}{K_a}\right)^2 - 4H_tG_t}}{2} \quad (\text{A.2.10})$$

Realistic (*i.e.* non-negative) values of concentrations are obtained only with *addition* of the square root expression. Equations A.2.10 and A.2.7 can be combined to give the observed chemical shift expressed as a function of the equilibrium constant, the total concentrations of host and guest, and the initial and final chemical shifts (Equation A.2.11). In this expression, the host concentration, guest concentration, and observed chemical shift are experimentally known. The concentration of the host can be maintained at a constant value by dissolving the guest in a stock solution of the host. This is of benefit only for weak association constants, where successive dilution will adversely affect the concentration of the complex or in situations where the NMR spectrum of the host is concentration-dependent. The use of constant host conditions avoids the participation of host aggregation in the shaping the binding isotherm. The chemical shift within the complex and the equilibrium constant are variables. The chemical shift of the host is an experimentally-known number only if the host exists purely as a monomer at the start of the titration. This is true for 9-ethyladenine in chloroform and methanol and δ_H is simply the observed chemical shift when no guest has been added.

$$\delta = \delta_H + (\delta_{HG} - \delta_H) \frac{\left(H_t + G_t + \frac{1}{K_a}\right) + \sqrt{\left(H_t + G_t + \frac{1}{K_a}\right)^2 - 4H_tG_t}}{2H_t} \quad (\text{A.2.11})$$

The equilibrium constant, K_a , and the chemical shift of the complex, δ_{HG} , can be determined by non-linear least squares regression curve-fitting of the experimental data to Equation A.2.11 using Systat.³⁵ This program uses either Simplex or Newton-Raphson algorithms to find the values of K_a and δ_{HG} which allow the best fit to Equation A.2.11.

The limitations of the NMR method for determination of association constants have been conscientiously summarized and elaborated by Wilcox.³ The important lesson from this is that the maximum amount of information in a NMR titration resides in data points from the range where the concentration of the complex is 20-80% of the maximum possible concentration of complex. That is, when the complex concentration is 20-80% of the lesser component in solution, host or guest. This imposes a practical determination limit of approximately 50000 M^{-1} , for the sensitivity of NMR.

A series of empirical studies on the effect of variation in concentration as a result of experimental error, evaporation, or impurity by Ghislain Deslongchamps³⁶ highlighted important experimental considerations that allow for accurate curve-fitting at the extreme edge of this limit. The conclusion from these studies was that experimental error in the concentration of the guest solution had a greater effect on K_a than error in the host concentration. However, the effect on the calculated association constant is approximately *three times* the

amount of variation in both host and guest concentration; that is, experimental error of $\pm 5\%$ in both guest and host concentration results in $\pm 15\%$ variation in the calculated association constant. In order to compensate for this, Equation A.2.11 was modified so that the initial concentration of the host solution could be considered a variable to be fit during regression analysis. That is,

$$H_i = \frac{V_i H_i}{V_i + V_a} \quad (\text{A.2.12})$$

where H_i is the variable initial host concentration and V_a is the volume of guest added, from the experimental data. The initial volume of the host solution, V_i , is a constant. The chemical shift is then expressed as a function of V_a and G_t with H_i , K_a , and δ_{HG} as variables to be determined (Equation A.2.13). By allowing the initial host concentration to vary, significantly improved fitting of experimental data can be achieved with concentration variations of *less than 5%*. This modification of the 1:1 binding isotherm provides an empirical scaling factor to compensate for poor data that can cause large error in cases where the binding constant is very large (*i.e.* $K_a \geq 10000 \text{ M}^{-1}$), such as the association constants determined in chloroform. While an extra parameter virtually guarantees improved curve-fitting, the experimental basis for the modification and its small magnitude justify judicious use of Equation A.2.13. This correction was only employed for the determination of the association of receptors **A5**, **A25**, **A26** in CDCl_3 .

$$\delta = \delta_H + (\delta_{HG} - \delta_H) \frac{\left(\frac{V_i H_i}{V_i + V_a} + G_t + \frac{1}{K_a} \right) + \sqrt{\left(\frac{V_i H_i}{V_i + V_a} + G_t + \frac{1}{K_a} \right)^2 - 4 \frac{V_i H_i}{V_i + V_a} G_t}}{2 \frac{V_i H_i}{V_i + V_a}} \quad (\text{A.2.13})$$

The association constants determined in CDCl_3 for the full complement of diimide receptors are listed in Table A.2.3. All of these association constants were determined by monitoring the ^1H NMR signals of a dilute solution of 9-ethyladenine (1-2 mM) during addition of the receptor (4-6 mM). The CH_2 quartet and C^8H were monitored. Concentration variations from 1-5% using Equation A.2.13 resulted in significantly improved agreement between the values determined for each signal. Because of the smaller size of 9-ethyladenine relative to the cleft, along with the greater mobility of the ethyl group, the CH_2 signals to remain sharp. In contrast, the aromatic signals of the adenine broaden during the titration. In all of the titrations in chloroform, a sharp break in the titration curve is seen when one equivalent of the cleft has been added to 9-ethyladenine. This is an indication of the 1:1 binding stoichiometry in the complex.

While C^8H can be monitored, the C^2H proton is usually not visible at all during the titration. This may be due to the differential effects that the environment of the 'bound' 9-ethyladenine has on the relaxation of magnetization.³⁰ The *spontaneous* decay of excited magnetization does not occur rapidly enough to occur at all on the NMR time scale – relaxation must be facilitated. The longitudinal relaxation of magnetization, T_1 , arises from the fluctuation of local magnetic fields of similar strength as the excited proton. This relaxation process is

aligned along the applied magnetic field and, by returning the excited proton to the ground state, is the process by which the FID is observed. Essentially, this results from the interaction of a proton with others -- on the same or nearby molecules -- through translation, rotation, conformational flexibility, or intramolecular dipole-dipole coupling. The C⁸H, CH₂, and CH₃ protons of 9-ethyladenine can all relax through dipole-dipole coupling, scalar (or spin-spin) relaxation, and spin rotation relaxation mechanisms, but the C²H proton is too isolated to participate efficiently through these pathways. When the translational and rotational movement of the adenine is reduced by complexation (the complex will have a much slower tumbling rate than the free molecules), both intermolecular dipole-dipole coupling and the chemical shift anisotropy relaxation are restricted. The effect of this is greater for C²H than for the other protons of 9-ethyladenine. In this case, then, *transverse* relaxation, T₂, will dominate. This relaxation process, which occurs perpendicular to the applied magnetic field, involves the dephasing of individual contributions to the overall magnetization. This results in a broadening of the signal due to that proton. An additional contribution to the broadening of the C²H proton may come from quadrupolar relaxation through the ¹⁴N of the imide to which it is hydrogen-bound. Quadrupolar relaxation is the dominant relaxation mechanism for ¹⁴N and acts to increase the transverse relaxation rate of coupled protons. This is the reason that NH protons typically have larger bandwidths than CH protons.

receptor	K _a (M ⁻¹)	ΔG ²⁹⁵ (kcal/mol)
A24^a	10 ⁶	-8
A23^a	10 ⁵	-7
A22^a	10 ⁶	-8
A21^a	10 ⁶	-8
A27^a	10 ⁶	-8
A26^b	53000	-6.4
A25^b	21000	-5.8
A5^b	50000	-6.3

Table A.2.3. Estimates of association constants of receptors with 9-ethyladenine in CDCl₃.
^acrude estimates only ^bdetermined by G. Deslongchamps using concentration correction

The core unit, **A5**, has an association constant of 50000 M⁻¹ (-6.3 kcal/mol).³⁶ This value is different from that reported earlier,²⁰ which was based on a titration following the imide proton as 9-ethyladenine was added to the receptor. The revised number is based on the 'reverse' titration protocol followed for the rest of the table. This association is the inherent binding affinity from the chelation of 9-ethyladenine along with stacking to carbazole. By altering the methyl ester to a benzyl ester (**A25**), the association constant decreases by 0.5 kcal/mol due to steric interaction within the complex between the ester and 9-ethyladenine.

Mutation of ester **A25** to amide **A26** leads to a 0.6 kcal/mol increase in binding energy in chloroform. As seen in the structural studies, this increase can be ascribed to a hydrogen bond between the amide and N³ of the adenine. Molecular modeling shows the hydrogen bond in a non-optimum geometry, which is also suggested

by the magnitude of the interaction (the strength of a base-pair hydrogen bond in chloroform has been estimated at between 1 and 1.5 kcal/mol^{12,14,37}). There are, in addition, two nearly free rotors that are restricted by forming this hydrogen bond – a negative contribution estimated to be worth ~1 kcal/mol each.³⁸

The fully elaborated receptors (**A21-A24**, **A27**) show extremely large binding affinities (Table A.2.3) for 9-ethyladenine in CDCl₃ – up to an *order of magnitude* greater than anything previously observed.^{4,15,39} The apparent binding constants were derived from curve fitting of 9-ethyladenine peak shifts upon addition of receptor and lie far outside of the range for which NMR titration is expected to provide accurate association constants.³ Remarkably, these numbers likely represent *lower* limits on the binding constants, since no attempt was made during the curve-fitting procedure to correct for aggregation of the receptors in solution.

The inaccuracy of the binding constants determined in CDCl₃ is too great to allow any comparison between them. For that reason, the association constants of the 'scorpion' receptors with 9-ethyladenine were determined in CD₃OD (Table A.2.4). Despite competition from the increased dipole moment and hydrogen-bonding capability of the solvent, these receptors demonstrate significant affinity for adenosine derivatives. All titrations were again performed using the 'reverse' titration protocol due to the flexibility inherent in these receptors. By monitoring purine signals, binding constants are derived from protons in similar environments in each of the different receptor-nucleoside complexes, thus reducing the effects of difference in local movement. Nevertheless, based on previous multiple determinations,¹⁴ the association constants are expected to be accurate only within 10% (± 0.1 kcal/mol). No concentration correction was incorporated in the determination of the binding constants.

receptor	K_a (M ⁻¹)	ΔG^{295} (kcal/mol)
A24	351	-3.4
A23	278	-3.3
A22	296	-3.3
A21	150	-2.9
A27^a	425	-3.5
A26^a	209	-3.1
A25^a	115	-2.8

Table A.2.4. Estimates of association constants of receptors with 9-ethyladenine in CD₃OD.

^adetermined by G. Deslongchamps

Despite the hydroxylic nature of the solvent, there is only modest propensity for aromatic stacking interactions in methanol. Titration of both 9-ethylguanine and 9-ethyladenine (1.5 mM) with 20 equivalents of the 'declawed' scorpion **A4** in CD₃OD resulted in less than 0.004 ppm upfield shifts of the purine aromatic signals.

The involvement of hydrogen bonding as a driving force for complexation was confirmed by a binding study performed with **A24** and 9-ethyladenine in CD₃OH. Binomial solvent suppression⁴⁰ was used to reduce the signal from the solvent hydroxyl protons. Upon addition of 2 equiv of 9-ethyladenine at -20 °C, the imide proton

signals of **A24** broadened and shifted 1.5 ppm downfield (10.7 ppm to 12.2 ppm). While broadening may be due to base-catalyzed exchange by adenine,⁴¹ the sizable downfield imide shifts provide convincing evidence for hydrogen bond interaction. This type of protocol is now used routinely in our laboratory to measure the association of water-soluble diimide clefts and adenosine derivatives in water.^{42,43}

Comparison of the binding affinities of receptors **A24** and **A27** and their incremental derivatives (**A21-A23**, **A25**, **A26**) allows an estimate of the contributions to binding in methanol by various structural elements (Table A.2.4). Energy differences of 0.1 kcal/mol are at the limit of experimental error, but changes greater than this are expected to be significant. A large proportion of the binding energy is derived from the simplest carbazole diimide motif (**A5**). The simultaneous Watson-Crick and Hoogsteen hydrogen-bonding 'chelation' of adenine provides the greatest stabilization with some contribution from aromatic stacking to the carbazole. The involvement of a hydrogen bond between the adenine N³ and amide NH² (**A26**) stabilizes complex formation by an additional 0.3 kcal/mol.

Elucidating the exact individual contribution by the upper amide (NH¹) and the aromatic cap is not possible. While the aromatic lid would be expected to provide stabilization through aromatic stacking interactions, the amide could be expected to participate in a hydrogen bond to N³, in a fashion similar to NH². Singly, though, each element appears to have a minor role in stabilizing the complex (0.1 kcal/mol, **A23** → **A24** and **A22** → **A24**). When the two are added together (**A21** → **A24** or **A26** → **A27**), significant stabilization (0.4-0.5 kcal/mol) results. The non-additivity of these 'mutations' complicates the analysis but suggests that both binding elements contribute to the binding energy of complexation, though not simultaneously. As a result, removal of one group allows the other to participate more fully, partially compensating for the loss.

There is a 0.2 kcal/mol increase in binding energy associated with the movement of the pendant phenyl ring one carbon away from the 3,6-diaminocarbazole (**A21** → **A26**). It is tempting to speculate that this is due to stabilization from edge-to-face aromatic interaction of the benzyl ring with the bound adenine. Theoretical studies suggest that this is more favorable for aromatic interactions than the planar stacked geometry.⁴⁴ Molecular modeling indicates that in the complex of 9-ethyladenine and **A21**, the geometry of the phenyl amide derivative precludes any interaromatic interaction when the amide NH²-N³ hydrogen bond is formed.

Examination of the magnitude of the binding energies in methanol (as compared to chloroform) allows an estimate of the strength of a hydrogen bond in this solvent. Both the overall affinity of **A25** for 9-ethyladenine and the increase in binding energy associated with the amide NH²-N³ hydrogen bond (**A26**) are reduced to one-half in methanol solution. This leads to the inference that the strength a single base-pair hydrogen bond contributes approximately 0.5-0.75 kcal/mol to binding in this solvent – half of the contribution in chloroform solution.

The selectivity of the carbazole-diimide motif was probed by titration of different nucleoside derivatives with **A24** in CD₃OD (Table A.2.5). The 0.3 kcal/mol increase in binding energy on complexation of 2'-

deoxyadenosine is likely due to additional hydrogen-bonding interactions between amide groups in the tail and the polar functionalities in the ribose group. The range of possible hydrogen-bonding patterns and conformational flexibility make this difficult to determine.

guest	K_a (M^{-1})	ΔG^{295} (kcal/mol)
9-ethyladenine	351	-3.4
9-ethylguanine	105	-2.7
2'-deoxyadenosine	577	-3.7
2'-deoxyguanosine	72	-2.5

Table A.2.5. Selectivity of **A24** for various nucleoside derivatives in CD_3OD .

While there is a distinct preference for adenine over guanine (3- to 8-fold), receptor **A24** shows surprisingly high affinity for guanosine derivatives. Molecular modeling studies indicate that the geometry of the imides may allow a three-hydrogen-bond interaction between carbazole bis-imide receptors and guanosine derivatives (Figure A.2.7). The reorientation of the imides and the subsequently different positioning of guanine within the cleft is likely responsible for the greater preference of the receptor for adenine in the nucleoside series versus the alkyl series (via altered hydrogen bonding to ribose). These studies suggest that for absolute selectivity, receptors must either have extreme rigidity or be specifically designed to select *against* the amino group on the guanine ring.

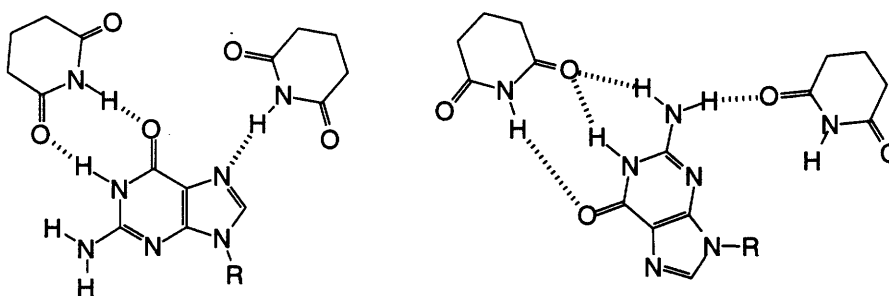


Figure A.2.7. Possible binding modes for guanosine derivatives.

3. DETERMINING HYDROGEN-BOND STRENGTH

The hydrogen bond is one of the fundamental forces for biomolecular recognition.⁴⁵ As such, there is tremendous interest in the contribution of a hydrogen bond to the stability of biological structures. Previous studies have attempted to determine this from model systems^{12,14} or from protein mutagenesis studies.^{24,46,47} The carbazole-diimide architecture can be exploited to provide a measure of this. Comparison of the binding affinity of diimide **A5** and the corresponding monoimide will allow an estimation of the enthalpy of a hydrogen bond.

Using the representation of Jencks^{48,49} as adapted by Williams,^{46,47} the free energy of an interaction can be broken down into contributions from *i*) intermolecular translational and rotational entropy, ΔG_{t+r} ; *ii*) entropy

loss from the restriction of rotors, ΔG_r ; *iii*) hydrophobic interactions, ΔG_h ; *iv*) the sum of the interactions between polar functional groups, $\Sigma\Delta G_p$; *v*) van der Waals interactions, ΔG_{vdW} , and *vi*) the change in conformational free energy between bound and free states. This is summarized in Equation A.3.1.

$$\Delta G = \Delta G_{t+r} + \Delta G_r + \Delta G_h + \sum \Delta G_p + \Delta G_{vdW} + \Delta G_{conf} \quad (\text{A.3.1})$$

For the comparison of diimide **A5** and monoimide **A30**, a rotor must be restricted for the second imide to bind to adenine. All other conformational, van der Waals, intermolecular, and hydrophobic terms are nearly identical between the two. In this case, then, the change in association free energy from **A30** to **A5** can be represented as:

$$\Delta\Delta G = \Delta\Delta G_r + \sum \Delta\Delta G_p \quad (\text{A.3.2})$$

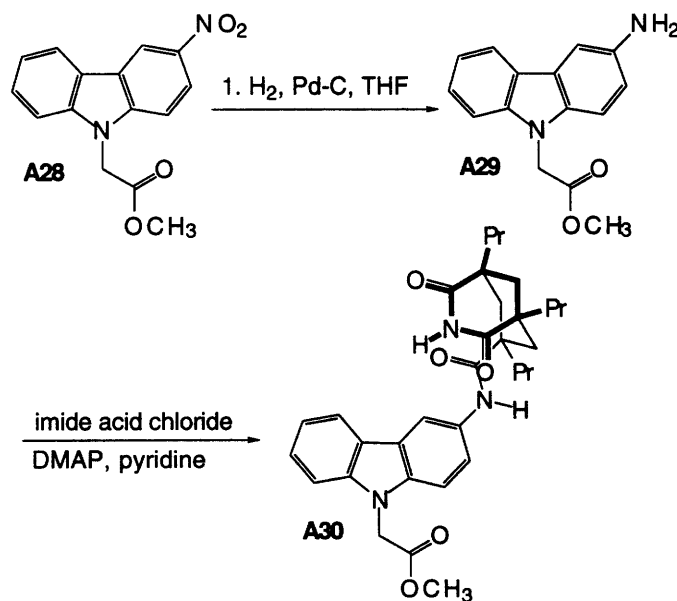
Or, more simply put:

$$\Delta G_{diimide} = \Delta G_{monoimide} + \Delta G_{base-pair} \quad (\text{A.3.3})$$

where,

$$\Delta G_{base-pair} = \Delta G_{H-bond} + \Delta G_{rotor} \quad (\text{A.3.4})$$

The monoimide receptor based on carbazole was required in order to examine these effects. Following alkylation of carbazole, **A7** could be quite cleanly *mono*-nitrated with copper(II) nitrate⁵⁰ to give 3-nitro-9-alkylcarbazole **A28**. While the second nitration step to 3,6-dinitrocarbazole is much slower than the first, the stoichiometry of reactions with traditional nitration reagents (*e.g.* HNO₃) are much more difficult to control than with this reagent. Additionally, the copper(II) nitrate reaction takes place at moderate temperatures (30 °C). Hydrogenation to the amine proceeded rapidly and cleanly with THF as solvent. This is in contrast to the original



Scheme A.3.1. Synthesis of carbazole monoimide.

conditions used in Section A.2 (HOAc/MeOH). Though the logic at that time was that protonation of the amine as it formed would increase its stabilization and reduce decomposition, methanol *itself* appears to contribute significantly to the decomposition of the amine. Condensation with the tripropyl derivative of Kemp's imide acid chloride²² gave monoimide receptor **A30** (Scheme A.3.1).

The affinity of the monoimide receptor for 9-ethyladenine was measured in CDCl₃ by monitoring the shift of the imide proton from 7.67-12.57 ppm during addition of 9-ethyladenine. The concentration of the imide was not kept constant through the course of the titration. Curve-fitting of the imide chemical shift to the 1:1 binding isotherm to give an association constant of 165 M⁻¹ (-3.0 kcal/mol). As described in the previous section, the diimide cleft **A5** has an association constant in CDCl₃ with 9-ethyladenine of 50000 M⁻¹ (-6.4 kcal/mol). This is the sum of the intermolecular interaction of 9-ethyladenine with a single imide, followed by intracomplex binding by the second. The 3.4 kcal/mol increase in binding energy of the diimide over the monoimide can be attributed solely to the hydrogen bonding of second imide within the complex. This energy difference is $\Delta G_{\text{base-pair}}$.

The bond that is restricted during binding of the second imide is the bond between carbazole C⁶H and the amide nitrogen. However, because of the steric and electronic nature of the amide group, this bond is somewhat less than a 'free' rotor. There are, in fact, only *two* low energy conformations expected for this torsion angle – 'above' and 'below' the carbazole plane. The only other non-bicyclic single bond in the second imide moiety is the cyclohexane-carbonyl bond which is so restricted by its proximity to the rigidly-held imide that it can be considered already fully 'restricted'.

The free energy change associated with the 'free' rotor can be estimated at:

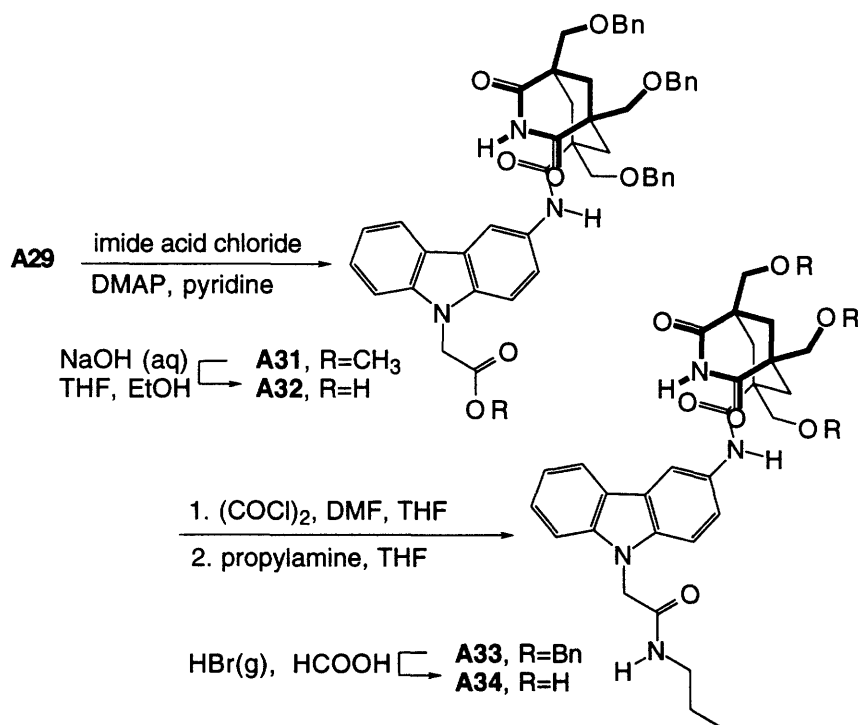
$$\Delta G_{\text{rotor}} \approx -T\Delta S_{\text{rotor}} = -RT \ln 1/2 = 0.4 \text{ kcal/mol} \quad (\text{A.3.5})$$

It is not surprising that this crude approximation for the restriction of a hindered rotation is somewhat less than the value of ~1 kcal/mol proposed by Williams for the restriction of a *free* rotor.³⁸

It follows, then, that $\Delta G_{\text{H-bond}}$ is 3 kcal/mol, 1.5 kcal/mol per hydrogen bond. It can be argued that by using the factorization approach, this value reflects the *enthalpy* of hydrogen-bonding. There are no surprises, though, from this determination. This value is very much similar to previous estimates determined by us and others.^{12,14,37}

There is considerably more controversy over the strength of a hydrogen-bond in water. Based on the binding of vancomycin antibiotic derivatives, Williams proposes an interaction energy of only 0.2-1 kcal/mol for an intermolecular hydrogen bond⁴⁶ whereas Fersht has concluded that a hydrogen-bond contributes 0.8 to 1.5 kcal/mol, based on analyses of protein and RNA mutations.^{17,24} It must be mentioned that the environment of protein interior is far different from aqueous solution, and may more accurately be expressed as resembling an organic solvent.

The same analysis as carried out on the chloroform-soluble system will be performed in *aqueous* milieu in collaboration with Yoko Kato. For the preparation of water-soluble imide receptors,⁴² amine **A29** was condensed with tris((benzyloxy)methyl) Kemp's imide acid chloride⁴² to give the benzyl-protected ester **A31**. The debenzylated ester was not sufficiently water-soluble for purposes of ¹H NMR titration, however propyl amide **A34** was found to dissolve easily, despite have two additional methylene groups. The large increase in water-solubility in the amide derivative serves as an important highlight of the hydrogen-bonding ability of amides over esters as a result of greater polarization between the nitrogen and carbonyl. An amide in this position is known to make a weak hydrogen bond to N³ of adenine in organic solvents (Section A.2.3), so this substitution must be maintained in the diimide to maintain the equivalence of the receptors. The synthesis of the diimide receptor and the measurement of binding affinity for these derivatives in aqueous solution will be performed by Yoko Kato.



Scheme A.3.2. Synthesis of water-soluble carbazole monoimide.

4 REINVESTIGATION OF NAPHTHYL-DIIMIDE RECEPTORS

The method for determining hydrogen-bond strength in water described in Section A.3 was initially applied to the receptor based on 2,7-diaminonaphthalene.⁵¹ It was found however, that in water there is almost *no* triplex hydrogen-bonding to adenine by the naphthalene-based receptor. The two imides appear to act almost completely independently as they each bind an adenine moiety. That is, the naphthalene-diimide functions to non-cooperatively bind two molecules of adenine. In contrast, the carbazole-based diimide receptor *has* been

shown to bind 9-ethyladenine and adenosine derivatives with triplex-like hydrogen bonding.⁴³

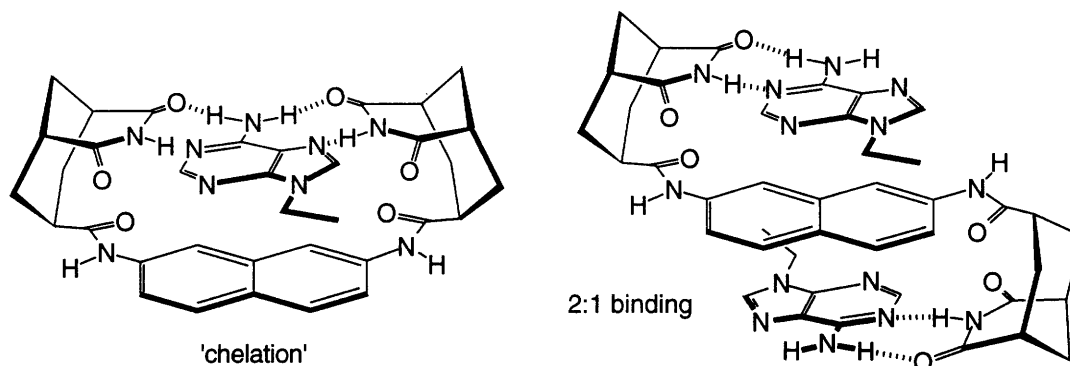


Figure A.4.1. The different binding modes of naphthalene-diimide receptor.

Though this suggests that the naphthalene framework is not optimum for simultaneous chelation of adenine, this type of behavior had not been initially observed for the 2,7-diaminonaphthalene-based receptors in organic solvents.^{14,15} While the early trimethyl Kemp's imide receptors were studied in mixtures of acetonitrile and chloroform,¹⁴ a more soluble version, **A36**, based on the tripropyl Kemp's triacid²² was calculated to have an association constant of 10^5 M^{-1} (-6.75 kcal/mol) in CDCl_3 by titration of the cleft with 9-ethyladenine. This titration was complicated by appreciable broadening of the imide signal and very little data could be taken for the early part of the titration curve. As a result, the majority of the data was obtained where the concentration of the complex was well over 80%.⁵² The poor quality of the data lowers our confidence in the curve-fitting results.

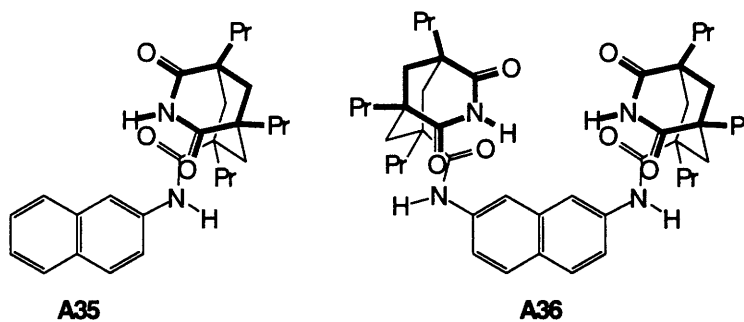


Figure A.4.2. Naphthyl-based imide clefts.

Receptors **A35** and **A36** were synthesized as previously described^{14,15} by coupling the commercially available amines with tripropyl Kemp's imide acid chloride.²² The association of **A36** and 9-ethyladenine was re-examined in CDCl_3 to ascertain the extent to which chelation occurs relative to 2:1 binding in this solvent system.

The association constant of monoimide **A35** and 9-ethyladenine was determined to be 188 M^{-1} (-3.06 kcal/mol) in CDCl_3 by monitoring the ^1H NMR chemical shift of the center of the CH_2 quartet as the receptor (49.8 mM) was added to a solution of 9-ethyladenine (maintained at 1.00 mM) in CDCl_3 . A dimerization constant of 13.8 M^{-1} for the imide was incorporated into the calculation using equations developed by Brian

Murray.⁵³ Namely, curve fitting (Systat³⁵) of the chemical shift of the naphthalene C¹H proton over a concentration range of 0.1-33 mM in CDCl₃ to Equation A.4.1 gave the dimerization constant, K_d, where δ_m is the chemical shift of the monomer, δ_d is the chemical shift of the dimer, and H_t is the concentration of the receptor. Both chemical shifts are asymptotes of the hyperbolic dimerization isotherm and must be calculated from the regression analysis. Incorporation of K_d as a constant into Equation A.4.2 allowed curve-fitting of the 1:1 titration curve between **A35** and 9-ethyladenine. The concentration of 9-ethyladenine, H_t, was a constant in this case. All other variables are defined as in Equation A.2.12 and A.4.1.

$$\delta = \delta_m + \left(\frac{\delta_d - \delta_m}{H_t} \right) \sqrt{\left(\left(H_t + \frac{1}{4K_d} \right) - \left(H_t + \frac{1}{K_d} \right)^2 - H_t^2 \right)} \quad (\text{A.4.1})$$

$$G_t = H_t \left(\frac{\delta - \delta_H}{\delta_{HG} - \delta_H} \right)^3 + \left(\frac{2K_d}{K_a^2} - G_t - 2H_t - \frac{1}{K_a} \right) \left(\frac{\delta - \delta_H}{\delta_{HG} - \delta_H} \right)^2 + \left(2G_t + H_t + \frac{1}{K_a} \right) \left(\frac{\delta - \delta_H}{\delta_{HG} - \delta_H} \right) \quad (\text{A.4.2})$$

In order to observe the complete saturation curve during titration of diimide **A36** and 9-ethyladenine, the 'reverse' titration protocol of section A.2 was performed – the diimide cleft was added to 9-ethyladenine and the ¹H NMR shift of the center of the CH₂ quartet was monitored. These titration conditions increase the influence of 2:1 binding on the saturation curve. The magnitude of the apparent binding constant is so much larger than the monoimide receptor that some amount of chelation *must* be occurring. In order for 2:1 binding, which is assumed to occur with the same affinity as the monoimide receptor, to compete with chelation, a large excess of 9-ethyladenine must be present (in the original titration procedure, an excess of 9-ethyladenine was present

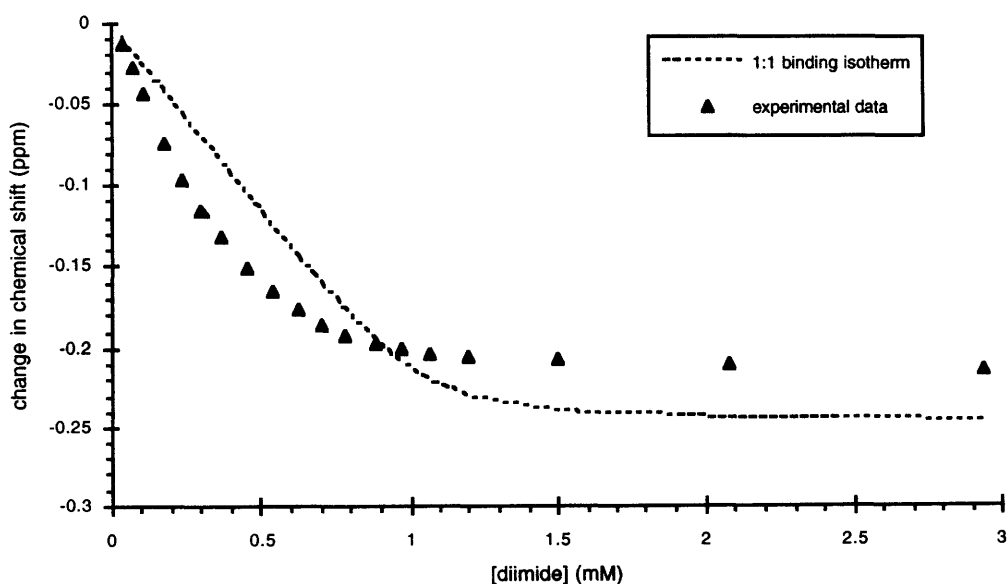


Figure A.4.3. Titration of 9-ethyladenine with naphthyl diimide does not fit 1:1 binding model.

only near the endpoint of the titration curve, where the imide signal is almost completely saturated). The data from this titration are shown in Figure A.4.3 along with the best fit from the 1:1 binding isotherm (Equation A.2.12). It is abundantly clear from this plot that the experimental data does *not* fit the purely 1:1 binding isotherm. In contrast, titration of 9-ethyladenine with carbazole diimide receptors gives no indication of this behavior.

The generalized binding scheme for mixed 1:1 (chelation) and 2:1 binding is outlined in Figure A.4.4. There are five equilibrium constants and three bound species in solution. No attempt is made to correct for the dimerization of either guest or host in chloroform. Even with judicious use of assumptions, this is a very difficult situation to describe mathematically. A solution for a similar binding situation has recently been described by Anslын.⁵⁴ The approach outlined here was solved by Belinda Tsao,⁵⁵ who also performed the curve-fitting itself.

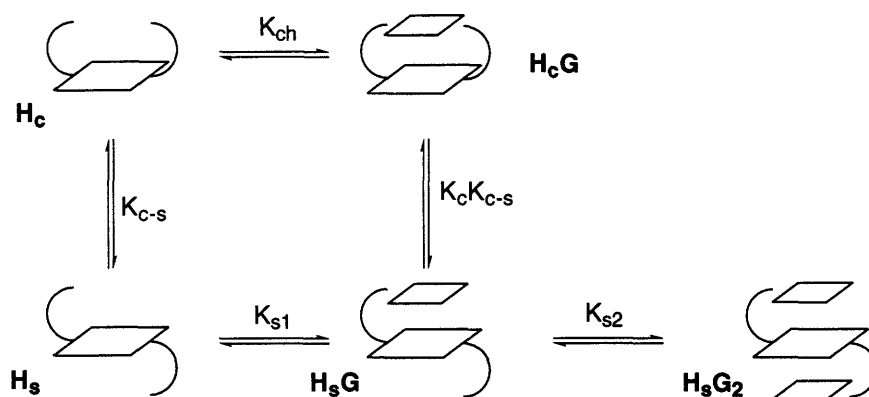


Figure A.4.4. Mixed 2:1, 1:1 chelation.

As the conformational and stoichiometric behavior of the receptor and target are dissimilar, the convention of the 1:1 binding isotherm that guest and host are determined by the titration protocol no longer applies. For this derivation, then, the host, H, is the receptor **A36** and the guest, G, is 9-ethyladenine. Given that H_c is the 'C-shaped' conformation of the host and H_s is the 'S-shaped' conformation, then the equilibrium constants depicted in Figure A.4.2 are defined in Equations A.4.3 to A.4.7.

$$K_{c-s} = \frac{[H_s]}{[H_c]} \quad (\text{A.4.3})$$

$$K_{ch} = \frac{[H_c G]}{[H_c][G]} \quad (\text{A.4.4})$$

$$K_{s1} = \frac{[H_s G]}{[H_s][G]} \quad (\text{A.4.5})$$

$$K_{s2} = \frac{[H_s G_2]}{[H_s G][G]} \quad (\text{A.4.6})$$

$$K_{c-s}K_c = \frac{[H_cG]}{[H_sG]} \quad (\text{A.4.7})$$

There are a number of assumptions which can be made to simplify this description. There is no interaction between the imides of **A36** that would be predicted to shift the equilibrium between the C-shaped and S-shaped conformations from unity. Accordingly, the value of K_{c-s} can be taken as 1, leading to the approximation:

$$[H_s] \approx [H_c] \equiv [H] \quad (\text{A.4.8})$$

where $[H]$ is the generalized concentration of free host.

The equilibrium between the chelated adenine and non-chelated adenine would be expected to lie *far* in favor of the chelated form so K_c can be assumed to be extremely small and the pathway corresponding to Equation A.4.7 can be neglected in further calculations. This assumption is implicit in the determination of hydrogen-bond strength of Section A.3. Also, since the monoimide binding of 9-ethyladenine is likely unaffected by previous association, K_{s1} and K_{s2} can be considered as essentially equivalent.

Manipulation of the equilibrium expressions gives the following equations for the variously bound species of receptor (Equations A.4.9 to A.4.11).

$$[H_cG] = K_{ch}[H][G] \quad (\text{A.4.9})$$

$$[H_sG] = K_s[H][G] \quad (\text{A.4.10})$$

$$[H_sG_2] = K_s^2[H][G]^2 \quad (\text{A.4.11})$$

The mass balance equations are, then:

$$H_t = [H] + [H_sG] + [H_{ch}G] + [H_sG_2] \quad (\text{A.4.12})$$

$$G_t = [G] + [H_sG] + [H_{ch}G] + 2[H_sG_2] \quad (\text{A.4.13})$$

The chemical shift of the 9-ethyladenine (maintained at constant concentration, G_t) is represented as a weighted average of the adenine chemical shifts in the unbound, singly-bound (chelated), singly-bound (monoimide), and doubly-bound complexes. That is,

$$\delta = \frac{[G]}{G_t} \delta_G + \frac{[H_{ch}G]}{G_t} \delta_{H_{ch}G} + \frac{[H_sG]}{G_t} \delta_{H_sG} + \frac{2[H_sG_2]}{G_t} \delta_{H_sG_2} \quad (\text{A.4.14})$$

By expressing the chemical shift as the chemical shift difference according to Equations A.4.15 to A.4.18, followed by simplification, the change in chemical shift can be represented as in Equation A.4.19.

$$\Delta = \delta_G - \delta \quad (\text{A.4.15})$$

$$\Delta_{H_{ch}G} = \delta_G - \delta_{H_{ch}G} \quad (\text{A.4.16})$$

$$\Delta_{H_sG} = \delta_G - \delta_{H_sG} \quad (\text{A.4.17})$$

$$\Delta_{H_sG_2} = \delta_G - \delta_{H_sG_2} \quad (\text{A.4.18})$$

$$\Delta = \frac{[H_{ch}G]}{G_t} \Delta_{H_{ch}G} + \frac{[H_sG]}{G_t} \Delta_{H_sG} + \frac{2[H_sG_2]}{G_t} \Delta_{H_sG_2} \quad (\text{A.4.19})$$

Incorporation of the equilibrium expressions (Equations A.4.4 to A.4.6) into the above equation gives Equation A.4.20, which expresses the chemical shift change as a function of the concentrations of free 9-ethyladenine and diimide.

$$\Delta = \frac{(K_s \Delta_{H_sG} + K_{ch} \Delta_{H_{ch}G} + 2K_s^2[G] \Delta_{H_sG_2})[H]}{1 + (K_s + K_{ch} + 2K_s^2[G])[H]} \quad (\text{A.4.20})$$

Since both K_s (188 M^{-1}) and Δ_{H_sG} (0.6077 ppm) are approximately known from the monoimide titration, this equation contains three variables. If the environment of the methylene group of 9-ethyladenine is considered to be in approximately the same environment in both H_sG and H_sG_2 , then

$$\Delta_{H_sG_2} \approx \Delta_{H_sG} \quad (\text{A.4.21})$$

This simplifies Equation A.4.20 to:

$$\Delta = \frac{(K_s \Delta_{H_sG} + K_{ch} \Delta_{H_{ch}G} + 2K_s^2[G] \Delta_{H_sG})[H]}{1 + (K_s + K_{ch} + 2K_s^2[G])[H]} \quad (\text{A.4.22})$$

There are now only two variables in the expression, K_{ch} and $\Delta_{H_{ch}G}$. However, a complication must now be addressed. This equation requires simultaneous knowledge of the *free* concentration of both host and guest. These values can be obtained from the mass balance equation A.4.13 with some manipulation. Multiplication of the multi-component terms by H_t/H_t and expanding the denominator gives an expression relating the free concentration of 9-ethyladenine to the equilibrium constants and total concentration of adenine and diimide at *each* point in the titration (A.4.23). With these values, the free host concentration can be calculated for each data point from the mass balance equation A.4.12 by substitution of equilibrium expressions A.4.9 to A.4.11 and knowledge of all binding constants.

$$G_t = [G] + \frac{((K_{ch} + K_s)[G] + K_s^2[G]^2)[H]_t}{1 + (K_{ch} + K_s)[G] + K_s^2[G]^2} \quad (\text{A.4.23})$$

The curve-fitting procedure, then, requires simultaneous regression analysis of two non-linear equations. Approximate values for the free host and guest concentrations are obtained by regression analysis of A.4.23 using an estimate for K_{ch} . The concentrations obtained are coupled with the experimental chemical shift data and used in the regression analysis of Equation A.4.22 to yield an 'improved' estimate of K_{ch} . This new estimate is, in turn, used for a second curve-fitting of Equation A.4.23. The iterative curve fitting of A.4.22 and A.4.23 is repeated until stable values for K_{ch} , $\Delta_{H_{ch}G}$ and the free host and guest concentrations are obtained.

The results of this iterative non-linear regression analysis are shown graphically in Figure A.4.5. The association constant is calculated to be 44000 M^{-1} (-6.3 kcal/mol) and Δ_{HchG} predicted to be 0.2202 ppm. The fit of the calculated data to the experimental numbers is very good ($\chi^2 = 3.3 \times 10^{-5}$).

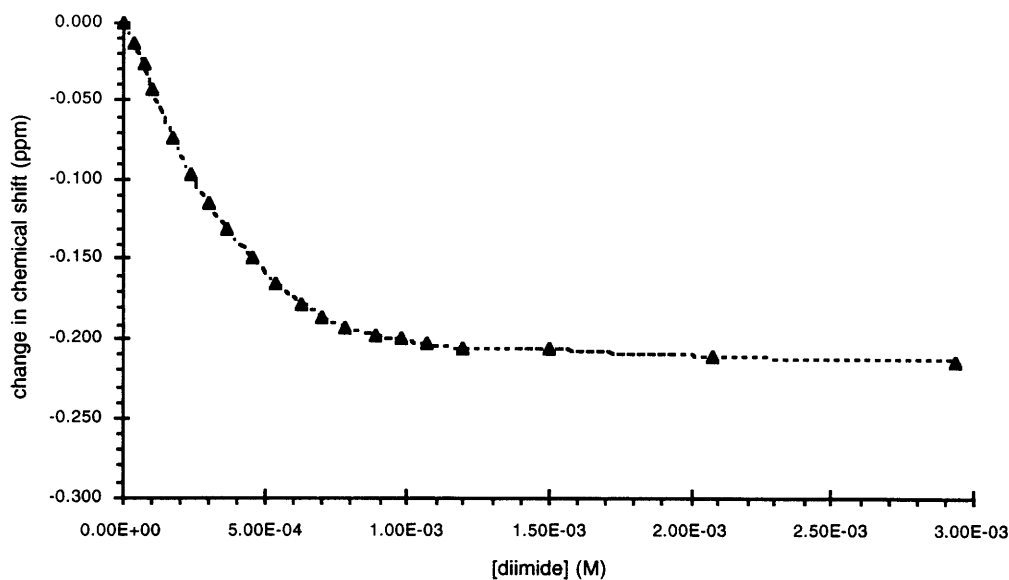


Figure A.4.5. Titration of 9-ethyladenine with diimide **A36** fits combined 1:1/2:1 model.

Despite the energetic barrier imposed by the intramolecular hydrogen bond between the two imides, the chelation binding affinity of **A5** (Section A.2) is approximately the same as the naphthalene-based receptor. This suggests that the *intrinsic* binding affinity of the carbazole-diimide architecture may be greater than the naphthalene-diimide by as much as the strength of a hydrogen bond, 1.5 kcal/mol.

While the angle imposed between the two imides on carbazole might be expected to have greater complementarity to adenine, an overlay of minimized conformations of the complexes shows that there is little difference in the positions of the imide groups as they bind to adenine (Figure A.4.6). In the structure of the carbazole-diimide complex with 9-ethyladenine, the Hoogsteen-bound imide is rotated in the plane of the aromatic to form a bifurcated bond to N^6H_2 . This type of interaction has previously been seen for simple monoimide amides, but not for the analogous imide esters, likely due to greater charge density on the amide carbonyl.⁵⁶ Though this would seem to lend only a minor increase in binding affinity, it may be a contribution the greater intrinsic affinity of the carbazole-based receptors.

As there would seem to be little difference between the hydrogen-bonding in the naphthalene- and carbazole-based diimide receptors, then the participation of 2:1 binding from the naphthalene system would appear to be linked to the aromatic stacking propensity of 2,7-diamidonaphthalene. From Figure A.4.7, which shows the superimposed aromatic region in both complexes, there is significantly more aromatic overlap in the complex with naphthalene, rather than carbazole. Additionally, the electron-rich nature of the carbazole group is not at all

like the electron-poor aromatic groups that typically intercalate into DNA. It is interesting to note that models based on offset aromatic stacked geometry,⁵⁷ and overlap of complementary aromatic point charges⁵⁸ would seem to predict that the carbazole moiety should stack more favorable with adenine than naphthalene.

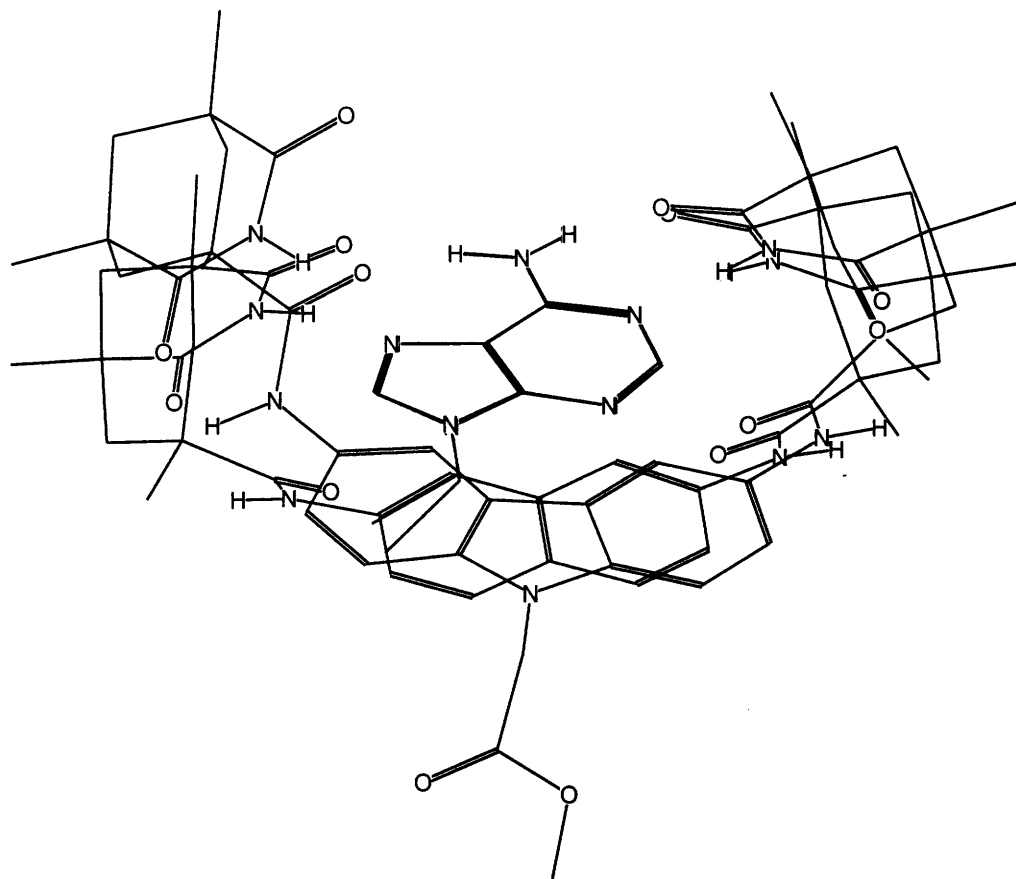


Figure A.4.6. Superimposed **A5** and **A36** complexes of 9-ethyladenine.

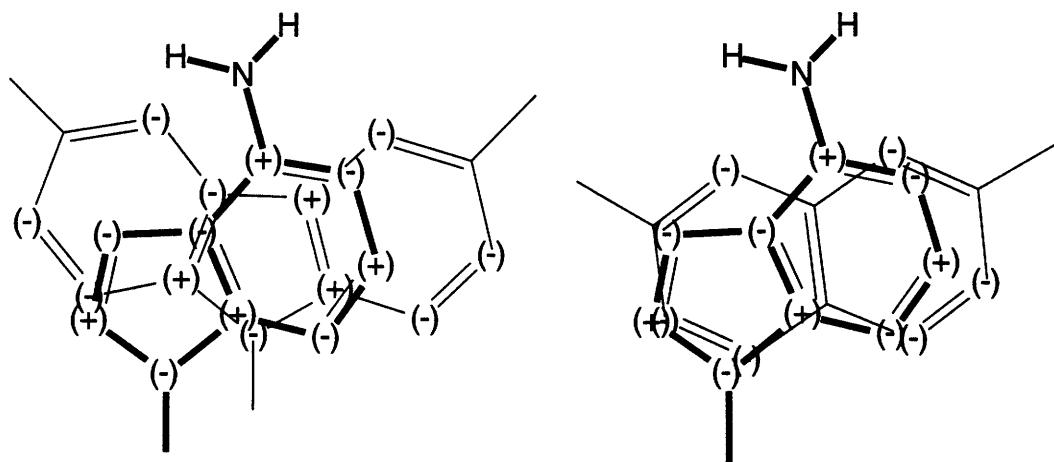


Figure A.4.7. Adenine superimposed on carbazole (left) and naphthalene (right) as in the complexes.

5. MONOIMIDE RECEPTORS

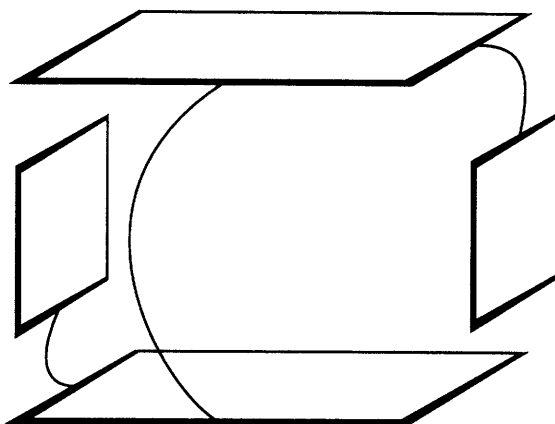


Figure A.5.1. Schematic of symmetric bis(imide)-adenine binding.

In Section A.2, a receptor was designed with an induced binding geometry. That is, the bound conformation of the receptor was defined by the binding event itself. The complex formed was a three-tiered aromatic 'sandwich', with aromatic stacking providing the driving force for the stacking of the third tier onto the complex. While in aqueous solution, the hydrophobic effect would provide a strong force for this in aqueous solution, there is very little driving force for aromatic stacking in organic solvents. It seemed appropriate, then, to design a similar receptor such that both top and bottom aromatic surfaces possessed similar affinity. This is outlined schematically in Figure A.5.1. This receptor is designed to be *structure-less* until the binding event occurs.

Here, each aromatic surface contains a single Kemp's imide, so that the driving force for complexation is hydrogen-bonding with concomitant aromatic stacking providing additional stabilization. This type of binding would also be predicted to be successful in water, where the hydrophobic effect provides the major contribution to binding while specificity for adenine contributed by hydrogen-bonding. Additionally, with a water-solubilizing tris(hydroxymethyl) Kemp's imide on each of the very hydrophobic carbazole moieties, the water-solubility might be enhanced relative to a two-imide/no-imide structure, as in Section A.2.

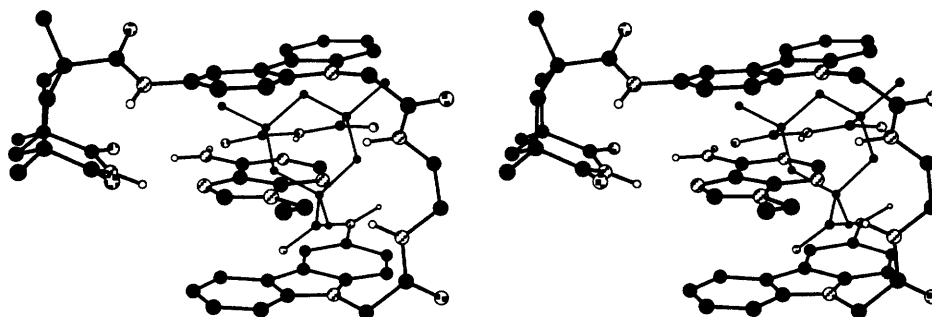
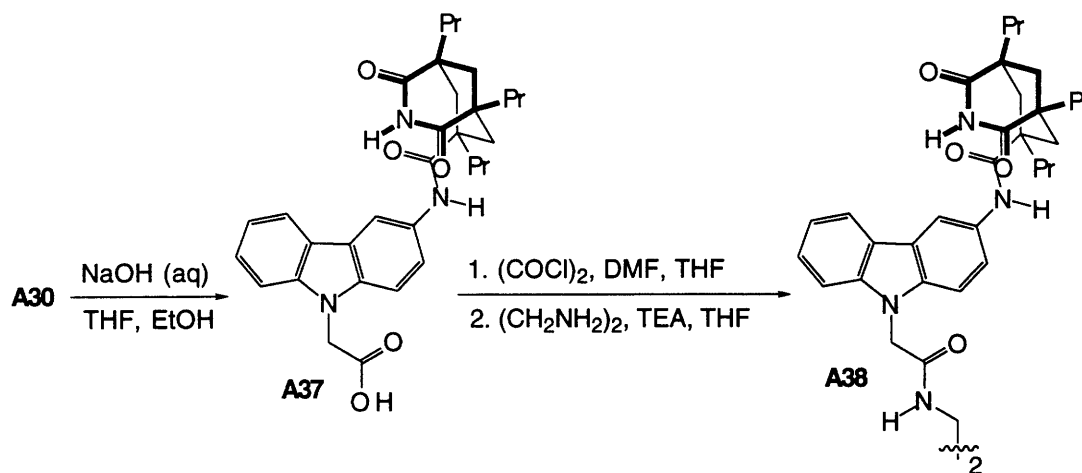


Figure A.5.2. Predicted geometry of mono-imide clefts induced by binding (stereoview).

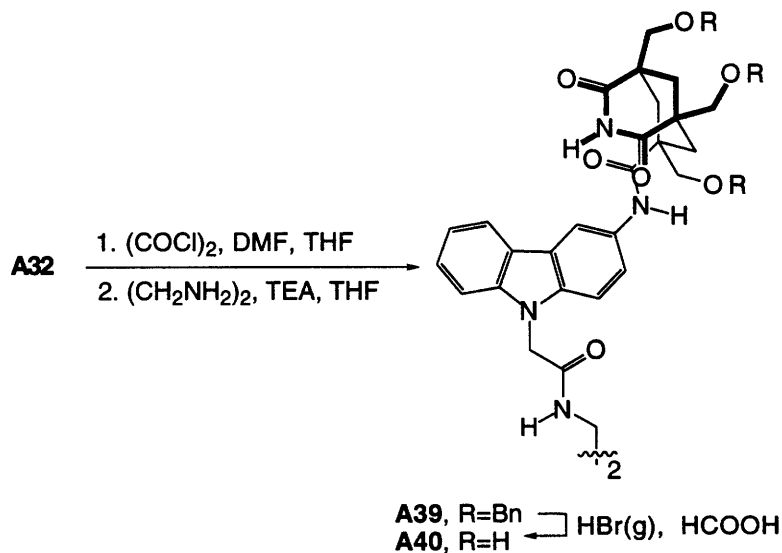
Molecular modeling was used to estimate the most favorable spacer for connecting the two aromatic surfaces

(Figure A.5.2). The lessons of Section A.2 led to the incorporation of an amide in the spacer to provide additional hydrogen bonding to adenine N³. Two units of carbazole monoimide **A30** are linked through an ethylenediamine chain to form the bis(monoimide) cleft **A38**. Base pairing can take place to the adenine, Watson-Crick from below and Hoogsteen from above, with coplanarity of all three aromatic surfaces. The backbone is arranged so that two partial hydrogen bonds are formed to adenine N³. The receptor has a C₂-like conformation in the complex.



Scheme A.5.1. Synthesis of induced-binding cleft.

The receptor was assembled from monoimide **A30** by hydrolysis, conversion to the acyl chloride, and condensation with ethylenediamine to form bis(imide) **A38** (Scheme A.5.1). The solubility of **A38** is extremely poor. Though 9-ethyladenine solubilized the cleft somewhat, the receptor could not be studied effectively. The flexibility of the spacer apparently allows an extensive intermolecular hydrogen-bonding network to form which is not adequately solvated by chloroform.



Scheme A.5.2. Synthesis of water-soluble induced-binding cleft.

Despite this, a water-soluble version of this cleft was synthesized with the hope that the greater solvating ability of water would break up the hydrogen-bond network to provide sufficient solubility and, as stated above, that the hydrophobic effect would provide adequate driving force for association. Acid **A32**, the tris((benzyloxy)methyl) version of the carbazole monoimide, was converted to the acid chloride and condensed with ethylenediamine. Deprotection with HBr(g) in HCOOH produced quantitative yield of the tris(hydroxymethyl) bis(imide) **A40**. Despite the breakup of the self-association, this cleft fell victim to the notoriously poor water solubility of carbazole and was insufficiently soluble in water (< 0.1 mM) for solution studies by proton NMR. Ultraviolet/visible spectroscopy is unsuitable for titrations because the accessible concentration range is too low to sample an adequate portion of the binding curve for the small binding constants expected in aqueous solvent.

6. EXPERIMENTAL SECTION

6.1 General

¹H NMR measurements were performed on Bruker AC-250 and AMX-300 and Varian XL-300, GE-300, UN-300, and VXR-500 spectrometers in solvents as indicated. Chemical shifts are reported as parts per million (δ) relative either to tetramethylsilane or to residual solvent peak. Melting points were obtained on an Electrothermal IA9100 digital melting point apparatus and are calibrated. IR spectra were recorded on a Mattson Cygnus 100 FT-IR spectrometer. High resolution mass spectra were obtained on a Finnegan MAT 8200 mass spectrometer. HPLC was performed on a Waters 600 Multisolvent Delivery System equipped with a Lambda Max Model 481 LC spectrophotometer using a Waters μ -Porasil column.

All commercially available compounds (Aldrich) were used without further purification unless otherwise indicated. *cis,cis*-1,3,5-Tripropylcyclohexane-1,3,5-tricarboxylic acid imide chloride was prepared according to Jeong *et al.*²² CD₃OD (99.8% D, Cambridge, Isotope Labs), CD₃OH (99% D, Cambridge Isotope Labs), CD₂Cl₂ (99.6% D, Aldrich Chemical), and DMSO-d₆ (99.9% D, MSD Isotopes) were used from freshly opened vials. CDCl₃ (99.8% D, MSD Isotopes) was stored over molecular sieves and passed through dry basic alumina just prior to use. Tetrahydrofuran and ether used in anhydrous conditions were distilled from sodium-benzophenone ketyl; dichloromethane and pyridine were distilled from calcium hydride; and N,N-dimethylformamide was dried over molecular sieves for several days. Thin layer chromatography was performed on Merck Silica 60 F-254 precoated TLC plates. Column chromatography was performed on Merck Silica Gel 60 (230-400 mesh) according to Still *et al.*⁵⁹ Glassware used for anhydrous conditions either was baked overnight at 150 °C, assembled hot, cooled under vacuum, and filled with argon; or was flamed dry under vacuum, cooled and filled with argon before use. Standard inert-atmosphere techniques were used for syringe and cannula transfers of liquids and solutions.

6.2 Synthesis

3,6-Dinitrocarbazole (**A1**)

was synthesized according to Grotta *et al*⁶⁰ and was purified by dissolving the crude mixture in hot 5:1 DMF/MEK (8 mL/g), filtering, cooling overnight at 3 °C, filtering and washing with cold 5:1 DMF/MEK, 5:3 DMF/MEK, and ether. The recovered solid was dried under vacuum (hot water bath) to yield the pure 3,6-dinitrocarbazole.

N-(Methoxycarbonylmethyl)-3,6-dinitrocarbazole (**A2**)

3,6-Dinitrocarbazole **A1** (10.12 g, 39 mmol) was stirred with dry K₂CO₃ (11 g, 2 equiv) in dry DMF (130 mL) at 40-60 °C for 1.5 h under argon. Methyl bromoacetate (8 mL, 2.2 equiv) was added to the dark red solution. The resulting cloudy orange mixture was stirred for 2 h at room temperature and for 1 h at 45-55 °C. The reaction was quenched with water (130 mL) and cooled on ice. The yellow product was filtered, washed with water, sucked dry overnight, and dried under vacuum (hot water bath) to yield **A2** (12.73 g, 99%) as a yellow solid: mp 314-315 °C; IR (KBr) 1750, 1340, 1310 cm⁻¹; ¹H NMR (300 MHz, DMSO-d₆) δ 9.538 (d, 2 H, J = 2.4 Hz), 8.445 (dd, 2 H, J = 9, 2.4 Hz), 7.924 (d, 2 H, J = 9 Hz), 5.639 (s, 2 H), 3.708 (s, 3 H); HRMS (EI) calcd for C₁₅H₁₁N₃O₆, 329.06478 found, 329.0645.

N-(Methoxycarbonylmethyl)-3,6-diaminocarbazole (**A3**)

A mixture of dinitro compound **A2** (2.41 g, 7.3 mmol), 10% Pd-C (520 mg, 22 wt %), MeOH (500 mL), and HOAc (50 mL) was hydrogenated at atmospheric pressure under a balloon for 5.5 h. The product mixture was filtered through a cake of Celite and the methanol removed by rotary evaporation under reduced pressure. Saturated NaHCO₃ was added to neutralize the acetic acid and the crude suspension was extracted with chloroform. The combined organic layers were washed with saturated NaHCO₃, and dried over anhydrous Na₂SO₄. Evaporation of the solvent under reduced pressure yielded 1.52 g (77%) of white solid **A3**, which darkens rapidly on exposure to light or air. The crude product is extremely sensitive to air and light and was used without further purification. ¹H NMR (product only) (300 MHz, CDCl₃) δ 7.314 (dd, 2 H, J = 2.1, 1.2 Hz), 7.079 (dd, 2 H, J = 8.7, 1.5 Hz), 6.865 (dd, 2 H, J = 8.4, 1.8 Hz), 4.879 (s, 2 H), 3.695 (s, 3 H), 3.596 (br s, 2 H).

N-((Methoxycarbonyl)methyl)-3,6-bis(pivaloylamino)carbazole (**A4**)

Diamine **A3** (770 mg, 2.86 mmol) in pyridine (50 mL) was treated with pivaloyl chloride (370 μL, 1.05 equiv) and stirred under argon at room temperature for 70 h. The solvent was removed by rotary evaporation under reduced pressure, and the residue was dissolved in chloroform, washed with 4% HCl, saturated NaHCO₃, and brine, and dried over MgSO₄. After concentration under reduced pressure, the crude product was purified by flash chromatography (3% MeOH/CHCl₃), trituration in cold methanol, and flash chromatography (10% EtOAc/CH₂Cl₂) to give 211 mg (17%) of an off-white powder: mp 242 °C dec; IR (KBr) 3425, 3297, 2963,

1732, 1657, 1529, 1479, 1198 cm^{-1} ; ^1H NMR (250 MHz, DMSO-d_6) δ 9.22 (s, 2 H), 8.34 (d, 2 H, $J = 1.4$ Hz), 7.55 (dd, 2 H, $J = 8.7, 1.6$ Hz), 7.44 (d, 2 H, $J = 8.8$ Hz), 5.29 (s, 2 H), 3.65 (s, 3 H), 1.26 (s, 18 H); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for $\text{C}_{25}\text{H}_{32}\text{N}_3\text{O}_4$ (M+H), 438.3293; found, 438.2391.

N-((Methoxycarbonyl)methyl)-3,6-bis(((cis,cis-2,4-dioxo-1,5,7-tripropyl-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino)carbazole (**A5**)

Diamine **A3** (0.4 g, 1.5 mmol) and *cis,cis*-1,3,5-tripropylcyclohexane-1,3,5-tricarboxylic acid imide chloride (1.055g, 2.06 equiv) were refluxed in dry pyridine under argon for 6.5 h. Solvent was removed by rotary evaporation, and the crude was suspended in chloroform (100 mL). The organic layer was washed with 10% HCl (100 mL) and brine (60 mL), and the layers were back extracted with chloroform (100 mL). The combined organic layers were concentrated by rotary evaporation to give a brown oil, which was triturated with cold methanol (6 mL) to give a beige powder, which was filtered and sucked dry. A second crop was recovered from the filtrate. Both crops of crystals and the final filtrate were purified by column chromatography to yield **A5** as a pale beige powder (520 mg, 39%): mp 227-229 °C; IR 3450, 3381, 3227, 3111, 2960, 2935, 2873, 1700, 1492, 1467, 1198 cm^{-1} ; ^1H NMR (250 MHz, DMSO-d_6) δ 10.40 (s, 2 H), 9.16 (s, 2 H), 8.11 (s, 2 H), 7.43 (dd, 4 H), 5.30 (s, 2 H), 3.63 (s, 3 H), 2.66 (d, 4 H, $J = 13.7$ Hz), 2.03 (d, 2 H, $J = 12.3$ Hz), 1.8-0.77 (m, 48 H); HRMS (EI) calcd for $\text{C}_{51}\text{H}_{69}\text{N}_5\text{O}_8$, 879.5146 found, 879.5146.

N-(Carboxymethyl)-3,6-bis(((cis,cis-2,4-dioxo-1,5,7-tripropyl-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino)carbazole (**A6**)

To a solution of ester **A5** (499 mg, 0.57 mmol) in THF (20 mL) was added 95% ethanol (60 mL), followed by 1 N NaOH (20 mL). The reaction solution was stirred under argon for 30 min and acidified to pH 1 with concentrated HCl. Solvent was removed by rotary evaporation and the crude material was triturated in water (20 mL) on ice, filtered, washed with cold water, sucked dry, and dried under vacuum to yield a pale beige solid (421.7 mg, 85%) which was used without further purification: ^1H NMR (300 MHz, DMSO-d_6) δ 10.39 (s, 2 H, imide), 9.16 (s, 2 H, amide), 8.11 (s, 2 H), 7.43 (dd, 4 H, $J = 14.4, 9$ Hz), 5.16 (s, 2 H), 2.67 (d, 4 H, $J = 13.2$ Hz), 2.03 (d, 2 H, $J = 12.6$ Hz), 1.6-0.7 (m, 48 H).

N-(Methoxycarbonylmethyl)carbazole (**A7**)

Carbazole (10.11 g, 60.5 mmol) was stirred with dry K_2CO_3 (22.24 g, 2.7 equiv) in dry DMF (100 mL) under argon for 45 min at 50 °C before methyl bromoacetate (17 mL, 3 equiv) was added portion-wise over 3.5 h. The reaction mixture was stirred at this temperature for 14 h and cooled. The reaction was quenched with cold water (200 mL) and triturated for 1 h on ice. The precipitate was filtered, washed with cold water (200 mL), and sucked dry. The product was dried under vacuum to yield white, crystalline **A7** (13.24 g, 91%): mp 97-98 °C; IR (KBr) 3054, 2936, 1731, 1457, 1328, 1267 cm^{-1} ; ^1H NMR (300 MHz, CDCl_3) δ 8.103 (d, 2 H, $J = 8.1$ Hz), 7.493 (t, 2 H, $J = 6.9$ Hz), 7.293 (m, 4 H), 5.02 (s, 2 H), 3.718 (s, 3 H); HRMS (EI) calcd for $\text{C}_{15}\text{H}_{13}\text{NO}_2$,

239.0946; found, 239.0944.

N-(2-Hydroxyethyl)carbazole (**A8**)

An anhydrous THF (30 mL) solution of the carbazole ester **A7** (5.3 g, 22 mmol) was cannulated over 15 min into an ice-cooled suspension of LiAlH₄ (0.94 g, 1.14 equiv) in THF (35 mL). The reaction mixture was stirred at room temperature for 1.75 h and was quenched with water (1 mL), 10% NaOH (1 mL), and water (3 mL).⁶¹ After the reaction mixture was stirred on ice for 10 min, the inorganic salts were removed by vacuum filtration and washed with ether (100 mL). The combined organic layers were washed with water to neutrality and with brine, and were dried over anhydrous MgSO₄. Solvent was removed by rotary evaporation under reduced pressure, and the residue was pumped free of solvent under vacuum. The crude brown oil recovered (4.48 g, 96%) was purified by column chromatography (25% ethyl acetate/hexanes) to yield white crystalline **A8** (3.06 g, 66%): mp 83-84 °C; IR (KBr) 3199, 3047, 1592, 1452, 1324 cm⁻¹; ¹H NMR (300 MHz, CDCl₃) δ 8.108 (dd, 2 H, J = 6.3, 1.2 Hz), 7.474 (d, 2 H, 3.9 Hz), 7.254 (m, 4 H), 4.486 (t, 2 H, J = 5.4 Hz), 4.062 (apparent q, 2 H, J = 4.5 Hz), 1.530 (s, OH); HRMS (EI) calcd for C₁₄H₁₃NO, 211.0997; found, 211.0995.

N-(2-Phthalimidoethyl)carbazole (**A9**)

DEAD (2 mL, 1.02 equiv) was added to a stirred solution of alcohol **A8** (2.62 g, 12.4 mmol), triphenylphosphine (3.29 g, 1.01 equiv), and phthalimide (1.88 g, 1.03 equiv) in anhydrous THF (150 mL) under argon. The reaction solution was stirred at room temperature for 22 h, refluxed for 4 h, and stirred at room temperature for an additional 44 h. Solvent was removed by rotary evaporation and the crude mixture pumped free of solvent under vacuum. The crude solid was triturated with ice-cold methanol (100 mL), and the off-white precipitate was filtered, washed with cold methanol, and dried to give crystalline **A9**: mp 143.1-143.6 °C; IR (KBr) 1772, 1710, 1598, 1384 cm⁻¹; ¹H NMR (300 MHz, CDCl₃) δ 8.082 (d, 2 H, J = 7.5 Hz), 7.816 (apparent dd, 2 H, J = 5.4, 3 Hz), 7.712 (apparent dd, 2 H, J = 5.7, 2.7 Hz), 7.534 (d, 2 H, J = 8.4 Hz), 7.437 (td, 2 H, J = 6.9, 1.5 Hz), 7.222 (td, 2 H, J = 7.5, 1.2 Hz), 4.595 (t, 2 H, 7.5 Hz), 4.124 (t, 2 H, 7.8 Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₂₂H₁₆N₂O₂, 340.1212; found, 340.1213.

N-(2-Aminoethyl)carbazole (**A10**)

The phthalimido compound **A9** (1.81 g, 5.3 mmol) was dissolved in THF (50 mL) and 95% ethanol (50 mL) with hydrazine monohydrate (530 μL, 2 equiv) and the resulting solution was refluxed for 18 h. Water (25 mL) was added, and the mixture was concentrated by rotary evaporation. Additional water (50 mL) and ether were added, and the mixture was brought to pH 14. The aqueous layer was extracted with ether, and the combined organic layers were washed with water and brine, and dried over anhydrous MgSO₄. The solvent was removed by rotary evaporation, and product was dried under vacuum to give off-white crystalline **A10** (1.16 g, 100%): mp 68-70 °C; IR (KBr) 3350, 1886, 1594, 1455, 1229 cm⁻¹; ¹H NMR (300 MHz, CDCl₃) δ 8.120 (dd, 2 H, J = 7.2, 1 Hz), 7.448 (d, 4 H, J = 3.9 Hz), 7.253 (m, 2 H), 4.413 (t, 2 H, J = 6.3 Hz), 3.229 (t, 2 H, J = 6 Hz), 1.152 (br s, 2

H); HRMS (EI) calcd for $C_{14}H_{14}N_2$, 210.1157; found, 210.1156.

3-(((Benzlyoxy)carbonyl)amino)benzoic acid (A11)

Benzyl chloroformate (1.15 mL, 1 equiv) in anhydrous THF (30 mL) was added to 3-aminobenzoic acid (1 g, 7.3 mmol) and dry pyridine (0.7 mL, 1.2 equiv) in anhydrous THF (200 mL) under argon. The reaction was stirred at room temperature for 2.5 h and solvent removed by rotary evaporation. The crude was triturated in water, acidified, filtered, sucked dry, and dried under vacuum to yield white crystalline product (1.945 g, 98%): mp 217.6-218.7 °C; IR (KBr) 3292, 1696, 1594, 1542, 1443, 1242 cm^{-1} ; 1H NMR (300 MHz, d_6 -acetone) δ 8.97 (s, 1 H), 8.307 (s, 1 H), 7.82 (d, 1 H, J = 9 Hz), 7.70 (d, 1 H, J = 8 Hz), 7.415 (m, 5 H), 5.189 (s, 2 H); HRMS (EI) calcd for $C_{15}H_{13}NO_4$, 271.0844; found, 271.0843.

Methyl 3-(((Benzlyoxy)carbonyl)amino)benzoate (A12)

Thionyl chloride (4 mL) was added to a suspension of acid **A11** (431.3 mg, 1.59 mmol) in dry methanol (40 mL). The mixture was stirred under a drying tube at room temperature for 5 h. Solvent was removed under reduced pressure, and the crude product was purified by flash chromatography (chloroform) to give white solid ester **A12** (349 mg, 77%): mp 124-124.6 °C; IR (KBr) 3344, 1736, 1697, 1547, 1211 cm^{-1} ; 1H NMR (250 MHz, $CDCl_3$) δ 7.97 (s, 1 H), 7.71-7.76 (m, 2 H), 7.36-7.42 (m, 6 H), 6.81 (br s, 1 H), 5.22 (s, 2 H), 3.90 (s, 3 H); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for $C_{16}H_{16}NO_4$ (M+H), 286.1079; found, 286.1077.

Methyl 3-Aminobenzoate (A13)

CBz-protected ester **A12** (230 mg, 0.81 mmol) was hydrogenated at balloon pressure for 10.5 h in THF (10 mL) and 95% ethanol (15 mL) with 10% palladium on carbon (40.5 mg, 18 wt %). Catalyst was removed by filtration through Celite and solvent removed by rotary evaporation under reduced pressure. The residue was dissolved in chloroform and dried over $MgSO_4$. Removal of solvent under reduced pressure followed by evacuation yields 98.8 mg (81%) of a pale-brown oil which was used without further purification: 1H NMR (250 MHz, $CDCl_3$) δ 7.43 (d, 1 H, J = 7.6 Hz), 7.35 (t, 1 H, J = 1.8 Hz), 7.22 (t, 1 H, J = 7.8 Hz), 6.86 (dd, 1 H, J = 7.8, 2.3 Hz), 3.89 (s, 3 H), 3.77 (br s, 2 H).

3-(((Benzlyoxy)carbonyl)amino)benzoyl Chloride (A14)

Acid **A11** (510.2 mg, 1.88 mmol) was refluxed in fresh $SOCl_2$ (15 mL) for 3 h. Thionyl chloride was removed by rotary evaporation with addition of benzene and dried under vacuum to give white crystalline **A14** (545 mg, 100%), which was used without further purification: IR ($CHCl_3$) 3430, 2253, 1793, 1745, 1741, 1606, 1593, 1530, 1489, 1439, 1431 cm^{-1} ; 1H NMR (300 MHz, $CDCl_3$) δ 8.083 (s, 1 H), 7.775 (d, 2 H, J = 7.2 Hz), 7.352 (m, 5 H), 7.113 (s, 1 H), 5.195 (s, 2 H).

N-Propyl-3-(((benzyloxy)carbonyl)amino)benzamide (**A15**)

Acid chloride **A14** (985 mg, 3.4 mmol) was dissolved in anhydrous CH₂Cl₂ (150 mL) under argon. Pyridine (280 μL, 1.02 equiv) and propylamine (560 μL, 2 equiv) were added. The solution was stirred at room temperature for 8 h and then washed with 5% (w/v) citric acid (50 mL), saturated NaHCO₃ (50 mL), and brine (50 mL). The organic layer was dried over anhydrous MgSO₄ and concentrated under reduced pressure. The crude product was recrystallized from ether/CH₂Cl₂ to give 636 mg (60%) of fluffy white crystals: mp 145.4-146 °C; IR (KBr) 3334, 3299, 2961, 1736, 1624, 1553 cm⁻¹; ¹H NMR (250 MHz, CDCl₃) δ 7.84 (s, 1 H), 7.61 (d, 1 H, J = 7.8 Hz), 7.45-7.33 (m, 7 H), 7.07 (br s, 1 H), 6.18 (br s, 1 H), 5.21 (s, 2 H), 3.39 (q, 2 H, J = 7.5 Hz), 1.609 (q, 2 H, J = 7.3 Hz), 0.964 (t, 3 H, J = 7.3 Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₁₈H₂₁N₂O₃ (M+H), 313.1552; found, 313.1550.

N-Propyl-3-aminobenzamide (**A16**)

The CBz-protected amine **A15** (361 mg, 1.16 mmol) was hydrogenolyzed using the same procedure as for **A13** to give colorless oil **A16** (203 mg, 99%), which darkened rapidly on exposure to light and air and was used without further purification: ¹H NMR (250 MHz, CDCl₃) δ 7.27 (s, 1 H), 7.16 (s, 1 H), 7.04 (br t, 1 H), 6.90 (m, 1 H), 4.15 (br s, 2 H), 3.49 (q, 2 H, J = 7.3 Hz), 1.08 (t, 3 H, J = 7.4 Hz).

N-(2-(9-*N*-Carbazolyl)ethyl)-3-(((benzyloxy)carbonyl)amino)benzoate (**A17**)

Acid chloride **A14** (515 mg, 1.78 mmol) and alcohol **A8** (375.8 mg, 1 equiv) were condensed using the procedure of **A15**. Recrystallization from chloroform gave 375 mg (45%) of **A17** as a white solid: mp 200-200.3 °C; IR (KBr) 3295, 3126, 1700, 1545, 1214 cm⁻¹; ¹H NMR (250 MHz, CDCl₃) δ 8.12 (d, 2 H, J = 7.9 Hz), 7.8 (d, 1 H, J = 8 Hz), 7.54-7.24 (m, 9 H), 6.5 (s, 1 H), 5.24 (s, 2 H), 4.73-4.67 (m, 4 H); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₂₉H₂₄N₂O₄, 464.1736; found, 464.1733.

N-(2-(9-*N*-Carbazolyl)ethyl)-3-aminobenzoate (**A18**)

The CBz-protected ester **A17** (199 mg, 0.428 mmol) was hydrogenolyzed using the procedure for **A13** to give 129.3 mg (91%) of **A18** as a colorless oil which was used without further purification: ¹H NMR (250 MHz, CDCl₃) δ 8.12 (d, 2 H, J = 7.8 Hz), 7.48-7.55 (m, 2 H), 7.22-7.37 (m, 5 H), 7.12 (t, 1 H, J = 7.8 Hz), 6.92 (t, 1 H, J = 1.9 Hz), 6.80 (dd, 1 H, J = 7, 2 Hz), 4.68 (m, 4 H), 3.62 (br s, 2 H).

N-(2-(9-*N*-Carbazolyl)ethyl)-3-(((benzyloxy)carbonyl)amino)benzamide (**A19**)

Amine **A10** (316.6 mg, 1 equiv) and acid chloride **A14** (435.8 mg, 1.5 mmol) were condensed using the same procedure as for **A15** to give off-white solid **A19** (686.3 mg, 99%), which was used without further purification: mp 207-211 °C dec; ¹H NMR (250 MHz, CDCl₃) δ 8.1 (d, 2 H, J = 6 Hz), 7.1-7.6 (m, 15 H), 5.200 (s, 2 H), 4.619 (t, 2 H, J = 6 Hz), 3.89 (apparent q, 2 H, J = 6 Hz).

N-(2-(9-*N*-Carbazolyl)ethyl)-3-aminobenzamide (**A20**)

The CBz-protected amide **A19** (586 mg, 1.26 mmol) was hydrogenolyzed using the procedure for **A13**. Following column chromatography (3.3% MeOH/CHCl₃), aromatic amine **A20** was recovered as an off-white foam (~50%), which resisted recrystallization attempts and was used without further purification: ¹H NMR (250 MHz, acetone-d₆) δ 8.129 (d, 2 H, J = 8 Hz), 7.82 (br m, 1 H), 7.613 (d, 2 H, J = 8 Hz), 7.419 (t, 2 H, J = 8 Hz), 7.189 (m, 2 H), 7.043 (t, 1 H, J = 7 Hz), 6.92 (d, 1 H, J = 8 Hz), 6.75 (dd, 1 H, J = 8, 2 Hz), 4.645 (t, 2 H, J = 6 Hz), 3.811 (apparent q, 2 H, J = 6 Hz).

Scorpion A21

Acid **A6** (253 mg, 0.29 mmol) was condensed with amine **A13** (67.2 mg, 1.52 equiv) using the method of scorpion **A24**. The product was purified by flash chromatography (2% MeOH/CHCl₃ and 2.3% MeOH/CHCl₃) followed by preparative HPLC (2% MeOH/CHCl₃) to remove residual **A5**. Pale beige **A21** (47.6 mg, 16%) was isolated: mp 193 °C dec; IR (KBr) 3380 (br), 2934, 2872, 1700, 1494, 1468, 1301, 1197 cm⁻¹; ¹H NMR (300 MHz, DMSO-d₆) δ 10.71 (s, 1 H, amide), 10.37 (s, 2 H, imide), 9.16 (s, 2 H, amide), 8.27 (d, 1 H, J = 1.8 Hz), 8.12 (s, 2 H), 7.82 (dd, 1 H, J = 7.8, 1.2 Hz), 7.64 (d, 1 H, J = 8.1 Hz), 7.45 (m, 5 H), 5.22 (s, 2 H), 3.81 (s, 3 H), 2.67 (d, 4 H, J = 13.5 Hz), 2.02 (d, 2 H, J = 12.6 Hz), 1.9-0.8 (m, 48 H); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₅₈H₇₅N₆O₉ (M+H), 999.5595; found, 999.5600.

Scorpion A22

Under argon, BOP-Cl (66 mg, 1.11 equiv) was added to acid **A6** (203 mg, 0.234 mmol) and triethylamine (33 μL, 1 equiv) in CH₂Cl₂ (20 mL) on ice. After 15 min stirring, the solution was allowed to come to room temperature and stirred for an additional 3 h. Amine **A16** (50 mg, 1.2 equiv) in CH₂Cl₂ (10 mL) was added with triethylamine (35 μL, 1.1 equiv) and the reaction mixture was stirred for 19 h at room temperature. The product solution was diluted with CH₂Cl₂ (100 mL), washed with 4% (v/v) HCl (50 mL), saturated NaHCO₃ (50 mL), and brine and was dried over MgSO₄. After removal of solvent under reduced pressure, the beige powder was purified by flash chromatography (EtOAc/CHCl₃ gradient) and preparative TLC (50% EtOAc/CHCl₃) to give pure **A22** as a beige powder (66 mg, 28%): mp 200 °C dec; IR (KBr) 3379, 2959, 2872, 1696, 1533, 1196 cm⁻¹; ¹H NMR (250 MHz, DMSO-d₆) δ 10.62 (s, 1 H, amide), 10.37 (s, 2 H, imide), 9.17 (s, 2 H, amide), 8.42 (t, 1 H, J = 7.3 Hz, amide), 8.12 (s, 2 H), 8.01 (s, 1 H), 7.73 (d, 1 H, J = 7.4 Hz), 7.51-7.33 (m, 6 H), 5.22 (s, 2 H), 3.17 (q, 2 H, J = 6.6 Hz), 2.67 (d, 4 H, J = 13.7 Hz), 2.02 (d, 2 H, J = 13.8 Hz), 1.9-0.78 (m, 48 H); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₆₀H₈₀N₇O₈ (M+H), 1026.6068; found, 1026.6055.

Scorpion A23

Acid **A6** (193.5 mg, 0.22 mmol), amine **A18** (83 mg, 1.13 equiv), N,N-bis(2-oxo-3-oxazolidinyl)phosphinic chloride (BOP-Cl, 64 mg, 1.13 equiv), and triethylamine (69 μL, 2.21 equiv) were mixed in anhydrous CH₂Cl₂

(20 mL), stirred under argon for 5 h, and refluxed for 16 h. The reaction solution was diluted with CH₂Cl₂ (50 mL) and washed with 5% (w/v) citric acid (25 mL), saturated NaHCO₃, and brine, and dried over MgSO₄. After removal of solvent by rotary evaporation under reduced pressure, the crude product was purified by flash chromatography (2.4% MeOH/CHCl₃) and preparative TLC (4% MeOH/CHCl₃) to give 33.2 mg (13%) of a beige powder: mp 180 °C dec; IR (KBr) 3378 (br), 2959, 2872, 1700, 1485, 1461, 1300, 1194 cm⁻¹; ¹H NMR (300 MHz, DMSO-d₆) δ 10.64 (s, 1 H, amide), 10.36 (s, 2 H, imide), 9.16 (s, 2 H, amide), 8.15 (m, 5 H), 7.7 (m, 1 H), 7.67 (d, 2 H, J = 8.1 Hz), 7.46 (m, 4 H), 7.35 (m, 4 H), 7.15 (t, 2 H, J = 6.7 Hz), 5.22 (s, 2 H), 4.79 (br s, 2 H), 4.60 (br s, 2 H), 2.67 (d, 4 H, J = 13.5 Hz), 2.02 (d, 2 H, J = 12.3 Hz), 1.9-0.75 (m, 48 H); HRMS (FAB in glycerol) calcd for C₇₁H₈₄N₇O₉ (M+H), 1178.6330; found, 1178.6328.

Scorpion A24

To a solution of acid **A6** (422 mg, 0.49 mmol) in anhydrous THF (40 mL) under argon was added CDI (88.9 mg, 1.12 equiv). The solution was stirred at room temperature for 1.6 h and amine **A20** (188.9 mg, 86% pure, 1 equiv) in anhydrous THF (10 mL) was added. The reaction solution was stirred at room temperature for 23 h and at 55 °C for 3.5 h. Solvent was removed by rotary evaporation. The residue was dissolved in chloroform, washed with 4% HCl, and brine and dried over anhydrous Na₂SO₄. Solvent was removed by rotary evaporation and the crude was purified by column chromatography (3% MeOH/CHCl₃ and 40% EtOAc/hexanes) to give **A24** as a pale brown foam (163.3 mg, 28%): mp 280 °C dec; IR (KBr) 3377 (br), 2959, 2932, 2872, 1700, 1485, 1465, 1195 cm⁻¹; ¹H NMR (250 MHz, DMSO-d₆) δ 10.607 (s, 1 H, amide), 10.377 (s, 2 H, imide), 9.169 (s, 2 H, amide), 8.624 (apparent t, 1 H, amide), 8.131 (s, 2 H), 8.115 (d, 2 H, J = 8.2 Hz), 8.036 (s, 1 H), 7.685 (m, 2 H), 7.578 (d, 2 H, J = 8.2 Hz), 7.454 (s, 4 H), 7.380 (t, 2 H, J = 7.3 Hz), 7.319 (d, 1 H, J = 4.9 Hz), 7.153 (t, 2 H, J = 7.4 Hz), 5.220 (s, 2 H), 4.536 (apparent t, 2 H), 3.606 (apparent q, 2 H), 2.667 (d, 4 H, J = 13.6 Hz), 2.023 (d, 2 H, J = 12.6 Hz), 0.772-1.9 (m, 48 H); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₇₁H₈₅N₈O₈ (M+H), 1177.6490; found, 1177.6482.

N-((Methoxycarbonyl)methyl)-3-nitrocarbazole (A28)

Alkyl carbazole **A7** (2.20 g, 9.194 mmol, purified by flash chromatography, 25-100% CHCl₃/hexanes) was added all at once to a solution of Cu(NO₃)₂ • 2.5H₂O (4.6 mmol, 1 equiv) in Ac₂O (10 mL) and AcOH (15 mL) at 32 °C under a drying tube. Additional AcOH (5 mL) was added to thin the suspension and the reaction was stirred for 5 h before being poured into water (300 mL). The mononitro product was isolated by filtration and drying to give 2.485 g (95%) of yellow solid which was used without further purification: ¹H NMR (300 MHz, acetone-d₆) δ 9.114 (d, 1 H, J=2.4 Hz), 8.401 (dd, 1 H, J=8.4, 0.6 Hz), 8.366 (dd, 1 H, J=9.0, 2.4 Hz), 7.738 (d, 1 H, J=9.0 Hz), 7.660 (d, 1 H, J=8.1 Hz), 7.589 (dd, 1 H, J=7.6, 1.1 Hz), 7.385 (dd, 1 H, J=7.3, 1.5 Hz), 5.431 (s, 2 H), 3.727 (s, 3 H).

N-((Methoxycarbonyl)methyl)-3-aminocarbazole (A29)

Nitrocarbazole **A28** (478 mg, 1.681 mmol) was dissolved in THF (50 mL) and hydrogenated at balloon pressure for 1.6 h with 10% Pd-C (218 mg, 46 wt%). The suspension was filtered through Celite and solvents removed by rotary evaporation and then under vacuum. The unstable solid was carried on immediately without further purification.

N-((Methoxycarbonyl)methyl)-3-(((cis,cis-2,4-dioxo-1,5,7-tripropyl-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino)carbazole (A30)

Amine **A29** (428 mg, 1.681 mmol) and tripropyl Kemp's acyl chloride²² (1.05 equiv) were refluxed with catalytic DMAP (0.034 equiv) in pyridine (11 mL) under a drying tube for 4.5 h and at room temperature for 16 h. After cooling, the reaction was taken up in EtOAc (150 mL) and washed with 1 N HCl (4 x 25 mL) and brine. The organic layer was dried over Na₂SO₄, concentrated, and purified by flash chromatography (2% MeOH/CH₂Cl₂) to give 817 mg (87%) of imide **A30**: IR (KBr) 3378, 2958, 1728, 1699, 1466, 1198, 746 cm⁻¹; ¹H NMR (500 MHz, CDCl₃) δ 8.240 (d, 1 H, J=2.0 Hz), 8.059 (d, 1 H, J=8.0 Hz), 7.737 (br s, 1 H, imide), 7.436 (t, 1 H, J=7.7 Hz), 7.390 (dd, 1 H, J=8.5, 1.5 Hz), 7.310 (s, 1 H, amide), 7.252 (d, 1 H, J=8.0 Hz), 7.227 (t, 1 H, J=7.7 Hz), 7.127 (d, 1 H, J=8.5 Hz), 4.905 (s, 2 H), 3.686 (s, 3 H), 2.651 (d, 2 H, J=14.0 Hz), 2.249 (d, 1 H, J=13.0 Hz), 2.003 (td, 2 H, J=13.0, 4.2 Hz), 1.55 (m, 2 H), 1.4-1.1 (m, 11 H), 0.95 (t, 6 H, J=7 Hz), 0.86 (t, 3 H, J=7 Hz); HRMS (EI) calcd for C₃₃H₄₁N₃O₅, 559.3046; found, 559.3041.

N-((Methoxycarbonyl)methyl)-3-(((cis,cis-2,4-dioxo-1,5,7-tris((benzyloxy)methyl)-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino)carbazole (A31)

In an analogous fashion as **A30**, tris((benzyloxy)methyl) Kemp's imide acid chloride⁴² (1.036 g, 1.799 mmol) was refluxed with aminocarbazole **A29** (1.5 equiv) in pyridine to yield 953 mg (67%) of the monoimide **A31** after gravity chromatography (10-15% EtOAc/CHCl₃), which was used without further purification: ¹H NMR (500 MHz, d₆-DMSO) δ 10.670 (s, 1 H, imide), 9.357 (s, 1 H, amide), 8.067 (d, 1 H, J=7.1 Hz), 7.5-7.2 (m, 21 H), 5.331 (s, 2 H), 4.511 (s, 2 H), 4.511 (s, 2 H), 4.506 (s, 2 H), 4.476 (s, 2 H), 3.772 (d, 2 H, J=9.0 Hz), 3.650 (s, 3 H), 3.458 (s, 2 H), 3.404 (d, 2 H, J=9.0 Hz), 2.545 (d, 2 H, J=14.4 Hz), 2.338 (d, 1 H, J=12.5 Hz), 1.608 (d, 2 H, J=13.7 Hz), 1.596 (d, 1 H, J=12.9 Hz).

N-((Carboxymethyl)-3-(((cis,cis-2,4-dioxo-1,5,7-tris((benzyloxy)methyl)-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino)carbazole (A32)

Imide **A31** (953 mg, 1.20 mmol) was hydrolyzed with 1 N NaOH (2 equiv) in THF (3.6 mL) and EtOH (2 mL) and stirring at room temperature for 6.5 h. Organic solvents were removed by rotary evaporation and the residue suspended in water (10 mL). After acidification with 10% HCl, the solid was isolated by filtration and dried under vacuum to give 831 mg (89%) of white powder which was used without further purification: HRMS

(EI) calcd for $C_{47}H_{45}N_3O_8$, 779.3207; found, 779.3210.

N-(((Propylamino)carbonyl)methyl)-3-(((cis,cis-2,4-dioxo-1,5,7-tris(benzyloxy)methyl)-3-azabicyclo [3.3.1]non-7-yl) carbonyl)amino]carbazole (A33)

Acid **A32** (102 mg, 0.131 mmol) was dissolved in anhydrous THF (4 mL) under argon. Catalytic DMF (<1 μ L) was added via a glass capillary, followed by oxalyl chloride (80 μ L, 7.0 equiv). The reaction was stirred at room temperature for 45 min, and solvent was removed by rotary evaporation. The residue was dried under vacuum and re-dissolved in anhydrous THF on ice under argon. Propylamine (150 μ L, 13.9 equiv) was added. The reaction was allowed to come to room temperature and stirred for a total of 53.5 h. Solvent was removed by rotary evaporation. The residue was taken up in chloroform (50 mL), washed with 1 N HCl (50 mL), and brine (25 mL), and dried over Na_2SO_4 . Evaporation of solvent followed by chromatography (30-50% EtOAc/ CH_2Cl_2) yielded **A33** (80 mg, 75%) as a white powder: IR (KBr) 3356, 1701, 1533, 1493, 1465, 1363, 1323, 1201, 1099, 746, 698 cm^{-1} ; 1H NMR (500 MHz, $DMSO-d_6$) δ 10.660 (s, 1 H, imide), 9.348 (s, 1 H, amide), 8.233 (br t, 1 H, amide), 8.063 (s, 4 H), 7.487 (d, 2 H, $J = 8.5$ Hz), 7.41-7.25 (m, 18 H), 7.185 (t, 1 H, $J = 7.3$ Hz), 4.969 (s, 2 H), 4.511 (s, 2 H), 4.505 (s, 2 H), 4.476 (s, 2 H), 3.772 (d, 2 H, $J = 9.0$ Hz), 3.461 (s, 2 H), 3.403 (d, 2 H, $J = 8.8$ Hz), 3.042 (q, 2 H, $J = 6.1$ Hz), 2.545 (d, 2 H, $J = 13.9$ Hz), 2.346 (d, 1 H, $J = 12.5$ Hz), 1.606 (d, 3 H, $J = 13.7$ Hz), 1.426 (q, 2 H, $J = 7.3$ Hz), 0.840 (t, 3 H, $J = 7.3$ Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for $C_{50}H_{53}N_4O_7$ (M+H), 821.3914; found, 821.3914.

N-(((Propylamino)carbonyl)methyl)-3-(((cis,cis-2,4-dioxo-1,5,7-tris(hydroxymethyl)-3-azabicyclo [3.3.1]non-7-yl) carbonyl)amino]carbazole (A34)

Benzyl-protected imide **A33** (72 mg, 87.7 μ mol) was dissolved in HCOOH (4.6 mL) and stirred, stoppered, on ice for 40 min. HBr (g) was bubbled through the solution for 9 min, followed by Ar (g) for 2 h. Solvent was removed by rotary evaporation. The residue was dissolved in MeOH with benzene and rotavapped dry. This was repeated twice and the residue dried *in vacuo* with a KOH trap. The waxy solid was dissolved in MeOH (1 mL) and precipitated by addition of ether (5 mL). The product was triturated briefly, filtered and dried under vacuum to give **A34** (40.5 mg, 84%) as a white solid: IR (KBr) 3326 (br), 1693, 1540, 1492, 1466, 1432, 1325, 1204, 1056, 804, 749 cm^{-1} ; 1H NMR (300 MHz, $DMSO-d_6$) δ 10.445 (s, 1 H, imide), 9.200 (s, 1 H, amide), 8.245 (br t, 1 H, amide), 8.115 (s, 1 H), 8.075 (d, 1 H, $J = 8.1$ Hz), 7.486 (d, 1 H, $J = 8.1$ Hz), 7.398 (m, 3 H), 7.177 (t, 1 H, $J = 7.5$ Hz), 4.969 (s, 2 H), 3.770 (d, 2 H, $J = 10.5$ Hz), 3.389 (s, 2 H), 3.318 (d, 2 H, $J = 10.2$ Hz), 3.040 (q, 2 H, $J = 6.5$ Hz), 2.375 (d, 2 H, $J = 14.1$ Hz), 2.205 (d, 1 H, $J = 12.6$ Hz), 1.487 (d, 2 H, $J = 14.7$ Hz), 1.425 (q, 2 H, $J = 6.9$ Hz), 1.328 (d, 1 H, $J = 12.9$ Hz), 0.839 (t, 3 H, $J = 7.5$ Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for $C_{29}H_{35}N_4O_7$ (M+H), 551.2506; found, 551.2510.

N-(Carboxymethyl)-3-(((cis,cis-2,4-dioxo-1,5,7-tripropyl-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino) carbazole (**A37**)

To a solution of ester **A30** (399 mg, 0.713 mmol) in THF (5 mL) and EtOH (10 mL) was added 1 N NaOH (1 mL, 1.4 equiv). TLC analysis indicated that the starting material had been completely consumed after 3 h of stirring at room temperature. At this time, solvents were removed by rotary evaporation and the residue dissolved in water (10 mL). Acidification to pH 1-2 with 10% HCl produced a beige precipitate, which was isolated by filtration and dried to give solid **A37** (365 mg, 94%) which was used without further purification.

1,2-Bis{*N*-[3-(((cis,cis-2,4-dioxo-1,5,7-tripropyl-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino) carbazolyl]methyl)carbonyl}amino}ethane (**A38**)

Tripropyl carbazole monoimide acid **A37** (285 mg, 0.522 mmol) was stirred in anhydrous THF (10 mL) with catalytic DMF under argon and (COCl)₂ (5.5 equiv) was added via syringe. The reaction was stirred for 1 h, and then solvent was removed, and the residue dried under vacuum for 2 h. The acyl chloride was then dissolved in anhydrous THF (10 mL) under argon and a solution of ethylenediamine (0.254 mmol) and TEA (2.152 mmol) in THF (5 mL) was added dropwise from an addition funnel. Additional THF (5 mL) was passed through the addition funnel and the reaction was stirred under argon for 2 h. The reaction was quenched with water (1 mL) and partitioned between CHCl₃ and 1 N HCl. The organic layer was washed with brine and dried over Na₂SO₄. After removal of solvent under reduced pressure, the residue was triturated in MeOH to isolate a pink solid which was passed through a short column (THF) to remove slow-eluting starting material. The residue was again triturated in MeOH to isolate a pink solid; 43 mg (15%): IR (KBr) 3352, 2958, 2932, 2871, 1696, 1661, 1537, 1492, 1465, 1430, 1321, 1193, 797, 746 cm⁻¹; ¹H NMR (500 MHz, DMSO-d₆) δ 10.398 (s, 2 H, imide), 9.199 (s, 2 H, amide), 8.316 (br t, 2 H, amide), 8.104 (s, 2 H), 8.096 (d, 2 H, J = 7.6 Hz), 7.463 (d, 2 H, J = 8.3 Hz), 7.38 (m, 6 H), 7.177 (t, 2 H, J = 7.8 Hz), 4.952 (s, 4 H), 3.161 (s, 4 H), 2.641 (d, 4 H, J = 13.9 Hz), 2.028 (d, 2 H, J = 12.1 Hz), 1.8-1.1 (m, 30 H), 0.866 (t, 12 H, J = 7.1 Hz), 0.797 (t, 6 H, J = 7.3 Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₆₆H₈₃N₈O₈ (M+H), 1115.6334; found, 1115.6322.

1,2-Bis{*N*-[3-(((cis,cis-2,4-dioxo-1,5,7-tris(benzyloxy)methyl)-3-azabicyclo[3.3.1]non-7-yl)carbonyl) amino]carbazolyl]methyl)carbonyl}amino}ethane (**A39**)

Oxalyl chloride (100 μL, 5.35 equiv) was added to a solution of acid **A32** (167 mg, 0.214 mmol) and catalytic DMF (<1 μL) in anhydrous THF (9 mL). The solution was stirred at room temperature for 1.5 h, solvent was removed by rotary evaporation, and the residue dried under vacuum. The acid chloride was dissolved in anhydrous THF (5 mL) and cooled on ice, under Ar. To this via cannula was added an ice-cooled solution of TEA (100 μL, 0.7175 mmol) and ethylenediamine (7.2 μL, 0.1077 mmol) in anhydrous THF (5 mL), followed by additional THF (4 mL) rinse. After stirring on ice for 15 min and at room temperature for 10.5 h, solvents were removed by rotary evaporation. The product was isolated by chromatography (5% MeOH/CH₂Cl₂) followed by

preparative TLC (3% MeOH/CH₂Cl₂, 2 mm plate, eluted twice) to give an off-white powder (75 mg, 44%): IR (KBr) 3331, 2922, 2859, 1702, 1590, 1529, 1493, 1465, 1232, 1198, 1099, 745, 698 cm⁻¹; ¹H NMR (300 MHz, DMSO-d₆) δ 10.662 (s, 2 H, imide), 9.3351 (s, 2 H, amide), 8.319 (s, 2 H, amide), 8.076 (s, 2 H), 8.063 (d, 2 H, J = 7.8 Hz), 7.5-7.1 (m, 40 H), 4.956 (s, 4 H), 4.504 (s, 8 H), 4.460 (s, 4 H), 3.768 (d, 4 H, J = 9.0 Hz), 3.450 (s, 4 H), 3.399 (d, 4 H, J = 9.3 Hz), 2.540 (d, 4 H, J = 15.3 Hz), 2.344 (d, 2 H, J = 13.2 Hz), 1.602 (d, 6 H, J = 14.7 Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₉₆H₉₅N₈O₁₄ (M+H), 1583.6968; found, 1583.6963.

1,2-Bis{N-[3-(((cis,cis-2,4-dioxo-1,5,7-tris(hydroxymethyl)-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino)carbazolyl)methyl}carbonyl}ethane (A40)

The benzyl-protected cleft (65 mg, 40.99 μmol) was dissolved in HCOOH, stoppered, and cooled on ice for 45 min. HBr(g) was bubbled through the solution for 7 min, followed by Ar(g) for 3 h. Solvent was removed by rotary evaporation and the residue pumped dry (KOH trap). The solid was dissolved in MeOH (4 mL) and precipitated with ether (6 mL). Following filtration, the solid was triturated in THF (4 mL) for 0.5 h, filtered, and dried under vacuum to give **A40** (39.1 mg, 92%) as a faintly yellow solid: IR (KBr) 3366, 2934, 2873, 1692, 1537, 1492, 1465, 1431, 1325, 1200, 1056, 806, 751 cm⁻¹; ¹H NMR (300 MHz, DMSO-d₆) δ 10.429 (s, 2 H, imide), 9.190 (s, 2 H, amide), 8.323 (br s, 2 H, amide), 8.129 (d, 2 H, J = 0.9 Hz), 8.076 (d, 2 H, J = 7.3 Hz), 7.464 (d, 2 H, J = 8.1 Hz), 7.393 (m, 6 H), 7.178 (t, 2 H, J = 7.3 Hz), 4.957 (s, 4 H), 4.0 (br s, 6 H, hydroxyl), 3.770 (d, 4 H, J = 10.5 Hz), 3.387 (s, 4 H), 3.319 (d, 4 H, J = 10.5 Hz), 3.172 (s, 4 H), 2.371 (d, 4 H, J = 13.2 Hz), 2.209 (d, 2 H, 12.9 Hz), 1.488 (d, 4 H, J = 13.8 Hz), 1.326 (d, 2 H, J = 13.2 Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₅₄H₅₉N₈O₁₄ (M+H), 1043.4151; found, 1043.4154.

6.3 Extractions

Extractions were performed by vigorously shaking 2 mL of a saturated aqueous solution of the nucleoside derivative with 2 mL of a 1 mg/mL solution of **A24** or **A27** in chloroform for 1 minute, drying the organic layer over Na₂SO₄, removing solvent under reduced pressure, and dissolving the residue in DMSO-d₆. Equivalents were calculated by integration of host and guest signals in the ¹H NMR spectrum. Extractions were run in duplicate. The absolute quantity of nucleoside extracted by the host was calculated by assuming 100% recovery of the receptor from the original extraction experiments. Blank extractions were performed similarly, but with no substrate in the chloroform. The absolute amount of nucleoside derivatives extracted was quantified using 3,6-dinitrocarbazole (1.14 mM in DMSO-d₆) as an internal standard.

6.4 Titrations

'Reverse' NMR titrations were performed at 250 MHz on a Bruker AC-250 instrument in a 5 mm NMR tube at ambient temperature (approximately 22 °C) by treating 400 or 500 μL of purine (1-2 mM) with aliquots of scorpion (4-6 mM) and monitoring the changes in the chemical shift of the substrate protons. For 'forward' titrations, scorpions (4 mM) were titrated with 9-ethyladenine (40 mM) in CD₃OD.

Nonlinear least squares regression curve fitting was used to determine binding constants, along with the final chemical shift of each of the monitored protons. In all cases, at least 50% of saturation was experimentally observed. The association constants given are the average of the binding constants obtained for each of the protons on the purine ring system. On the basis of our previous results,¹⁴ the calculated values are expected to be no more accurate than within 10% of the actual value. For chloroform titrations, where association constants are extremely large, a concentration correction was introduced into the regression analysis. The variations in concentration so determined were less than 1%, but resulted in a significantly better fit of the data.

6.5 NMR Spectroscopy

Processing and analysis of spectra were done on instrument. Experiments performed on Bruker instruments used shifted sinebell FID weighting while those done on Varian instruments employed a Gaussian multiplication of FID's. Except for the Bruker AC-250, which used time proportional phase increments (TPPI),⁶² phase cycling was accomplished through the use of hypercomplex data.⁶³

Longitudinal relaxation times (T_1) were measured using the inversion-recovery method as implemented in the Varian software (VNMR version 3.1).

Proton resonance assignments of **A24** were made in DMSO- d_6 from DQF-COSY^{64,65} (AC-250, 256 x 512 data points, 292 Hz spectral width, zero-filled to 512 x 512) and NOESY^{29,66,67} (UN-300, mixing time of 1 s, 512 x 4K data points, 3919 Hz spectral width, zero-filled to 2K x 4K).

Assignment of resonances in the proton NMR spectrum of the one-to-one complex of **A24** and 9-ethyladenine in $CDCl_3$ were based on a DQF-COSY experiment on UN-300 (150 x 512 data points, 1220 Hz, zero-filled to 512 x 512).

Low temperature studies were performed with **A24** in $CDCl_3$ on a Varian GE-300 (2.1 and 12.2 mM in chloroform) and with the one-to-one complex of **A24** and 9-ethyladenine on a Varian VXR-500 in $CDCl_3$ (5.5 mM) and in CD_3OH (2 mM). Hydroxyl protons of CD_3OH were suppressed using 1:3:3:1 binomial solvent suppression.^{40,68}

Peak assignments and through-space correlation of the one-to-one complex of **A27** and 9-ethyladenine in CD_2Cl_2 were made on the basis of TOCSY (mixing time of 30 ms, 512 x 4K data points, 4500 Hz, zero-filled to 2K x 4K) and NOESY (mixing time of 500 ms, 384 x 4K data points, 4425 Hz, zero-filled to 2K x 4K) on UN-300.

Peak assignments and through-space correlations of the one-to-one complex of **A24** and 9-ethyladenine in CD_2Cl_2 were made on the basis of TOCSY^{28,69} (mixing time of 25 ms, 512 x 4K data points, 4200 Hz in ω_1 and 5434 Hz in ω_2 , zero-filled to 2K x 4K) and NOESY (mixing times of 250 ms and 500 ms, 256 x 4K data points, 4200 Hz in ω_1 and 5434 Hz in ω_2 , zero-filled to a 2K x 4K) experiments on AMX-300. These experiments were performed at the Universitat de les Illes Balears by Prof. Pablo Ballester of U.I.B. and Dr. Alain Pagelot of Sadis Bruker Spectrospin SA.

Low temperature (-50 °C) peak assignments of the complex of **A24** and 9-ethyladenine in CDCl₃ were made on the basis of TOCSY (mixing time of 30 ms, 2K x 300 data points, 1800 Hz, zero-filled to 2K x 1K) and NOESY (mixing time of 500 ms, 4K x 666 data points, 8000 Hz, zero-filled to 4K x 2K) experiments on a Varian VXR-500 spectrometer equipped with a 5-mm ¹H/¹⁹F inverse detection probe and an Oxford VTC-4 temperature controller.

6.6 Molecular Modeling

All molecular modeling was performed on Silicon Graphics Personal Iris workstations (4D25G+ or 4D30G+) using MacroModel 3.5X.⁷⁰ Stochastic dynamics⁷¹ simulation of **A24** and 9-ethyladenine in chloroform was run in the MacroModel version of the AMBER⁷² force field (AMBER*) at 300K with GB/SA continuum solvation,⁷³ extended non-bonded distance cutoffs, constrained bond length (SHAKE), and a 1.5-fs timestep using a minimized conformation of the complex as the starting point. After a 5.0-ps initialization period, structures were sampled every 0.2 ps for 60 ps. The 300 structures generated during the stochastic dynamics simulation were minimized using the MULT routine in MacroModel to yield 31 unique conformations, using 0.4 Å as the maximum allowed deviation for comparison of all heavy atoms.

B. NUCLEOTIDE RECOGNITION

1. INTRODUCTION

As the keystone of DNA backbone structure, the chemistry of the phosphate group dominates transcription, replication, translation, and energy transfer processes. Molecular recognition studies targeted for phosphates or nucleotides have a number of ultimate goals, including the sequence-specific recognition of DNA and transport of RNA or various nucleoside-based drugs across cell membranes. A receptor for the phosphate diester must provide both hydrogen-bond complementarity to the two phosphoryl groups and charge complementarity to the single negative charge. The molecular shapes that have been employed for this include protonated aza crown ethers,^{7,74} bicyclic guanidinium cations,^{75,76} expanded porphyrins,⁷⁷ bis(acylguanidiniums),⁷⁸ and bis(alkylguanidiniums).⁵⁴

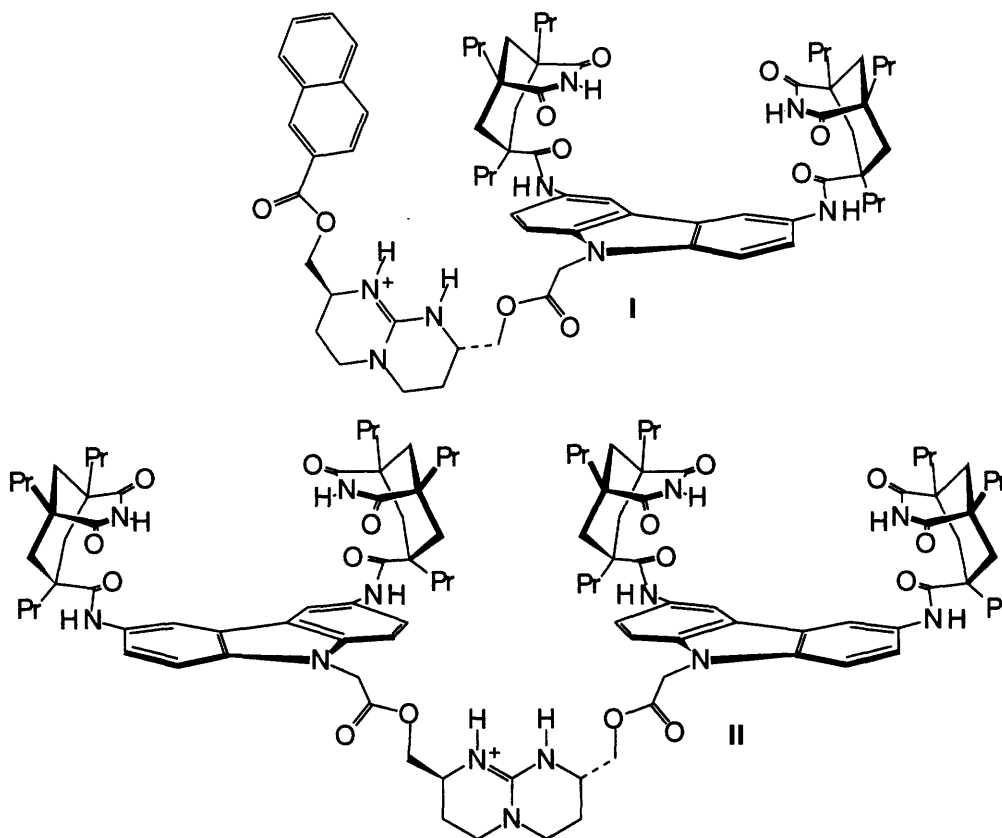


Figure B.1.1. First-generation nucleotide transport agents.

When these phosphate recognition elements are coupled with subunits for recognition of nucleobases, the molecular recognition of nucleotides is possible. Aromatic groups based upon DNA intercalators are often used for this purpose,^{7,79} but hydrogen-bonding complementarity has also been employed.⁸ Nucleotide receptors have been used for transport of nucleotides across liquid membranes^{77,80-82} and for the catalysis of

phosphodiester formation⁷⁴ and cleavage.^{7,83-85}

The carbazole-diimide cleft can be combined with a bicyclic guanidinium⁷⁵ to yield a nucleotide receptor specific for adenosine nucleotides. Specifically, esterification of diimide **A5** with a guanidinium alcohol gave receptor **I** (Figure B.1.1) which was found to extract adenosine nucleotides from aqueous solution into chloroform.¹⁹ The span of the recognition could be extended by linking two carbazole-diimide receptors to the guanidinium diol (**II**, Figure B.1.1) yielding a receptor specific for adenosine dinucleotides.¹⁸ These receptors were developed in collaboration with Javier de Mendoza of the Universidad Autonoma de Madrid, Spain. The structures of the complexes of these two receptors with their targets were determined by ¹H NMR ROESY experiments. In both complexes, the nucleotide is bound by the receptor through charge-charge interactions at the guanidinium-phosphate salt-bridge and hydrogen-bonding in the diimide cleft.

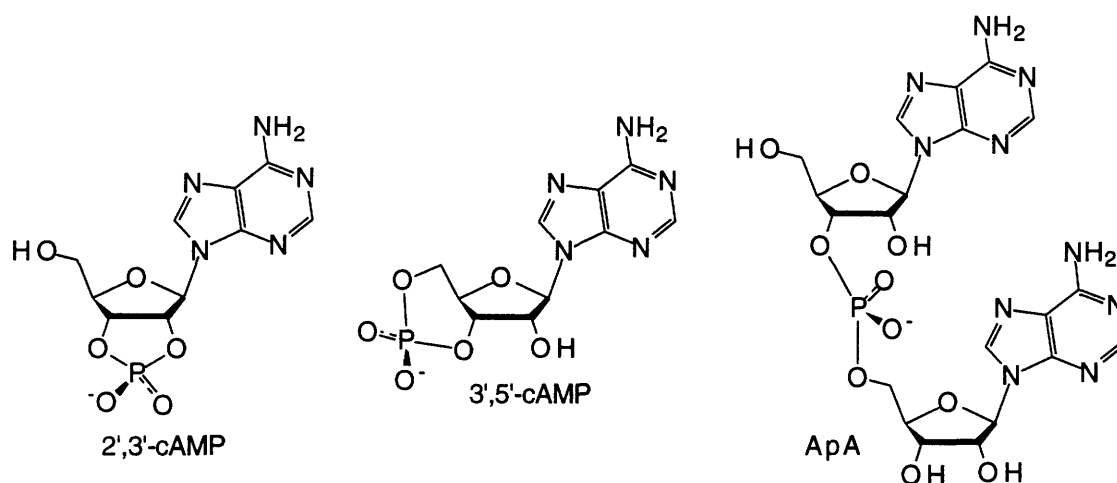


Figure B.1.2. Adenosine nucleotide targets.

As the first rung in the ladder of adenosine-specific nucleotide recognition, receptors **I** and **II** are quite successful. However, tuning of the architecture of these receptors will allow greater control of the binding process (*e.g.* to gain greater specificity). The linkage between carbazole and guanidinium contains a significant amount of flexibility; there are four free rotors between the two recognition motifs. This allows the guanidinium moiety to explore a wide range of conformation space with concomitant entropy losses on binding caused by the freezing out of these rotors. This conformational flexibility, when extended to longer nucleotide systems, can complicate the binding process and examination of it.

One of the ways in which this conformational flexibility can be reduced is to replace the methylene spacer of **I** and **II** with a phenyl spacer. This would serve to reduce the space swept out by these rotors from a cone to a cylinder. This greatly reduces the amount of space explored by the bicyclic guanidinium and, as a consequence, the entropy lost by restricting this group during binding. The geometry of a *para*-substituted phenyl group allows purely 'in-line' displacement of guanidinium from the carbazole; the bond angles of a spacer based on tetrahedral carbon will not allow this. In addition, the phenyl group provides a slightly expanded space for recognition, which should allow more freedom for the adenine to bind completely within the diimide cleft. The

more rigid geometry of a receptor with a phenyl group could confer an advantage in both mononucleotide binding, in the form of specificity, and receptors for longer nucleotides, in the form of better preorganization of the binding geometry. In addition to the structural advantages to a phenyl spacer, there will also be greater chemical stability to esters of this type. The alkyl esters of **I** and **II** have been found to be quite labile in hydroxylic or polar organic media; significant hydrolysis of this spacer occurs within minutes in methanol.

2. SELECTIVE CYCLIC MONOPHOSPHATE BINDING

Preliminary molecular modeling of a simple mononucleotide receptor based upon a phenyl-substituted carbazole yielded an exciting prediction. This simple receptor may be able to selectively bind 2',3'-cAMP over 3',5'-cAMP. The origin of this behavior would be a salt-bridge hydrogen-bond that could only be formed in the complex with 2',3'-cAMP. With the rigidity imparted by receptor and the substrates, this salt-bridge interaction should not form in the complex with 3',5'-cAMP.

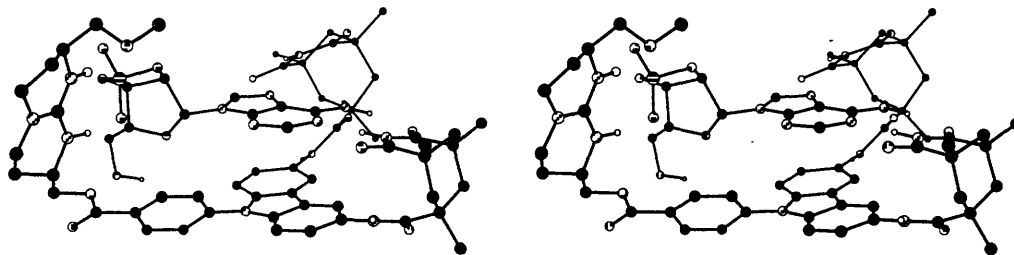


Figure B.2.1. Predicted structure of **B9** complexed with 2',3'-cAMP (stereoview).

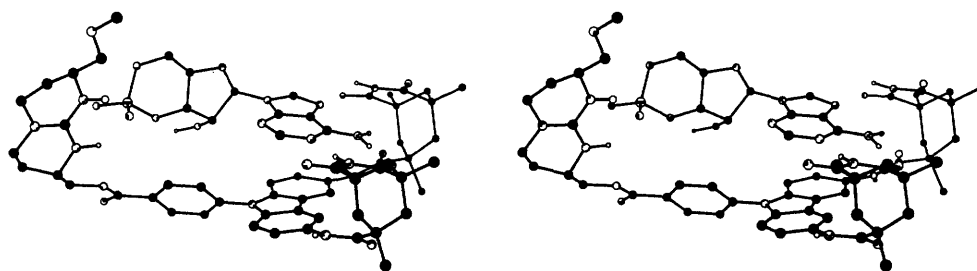
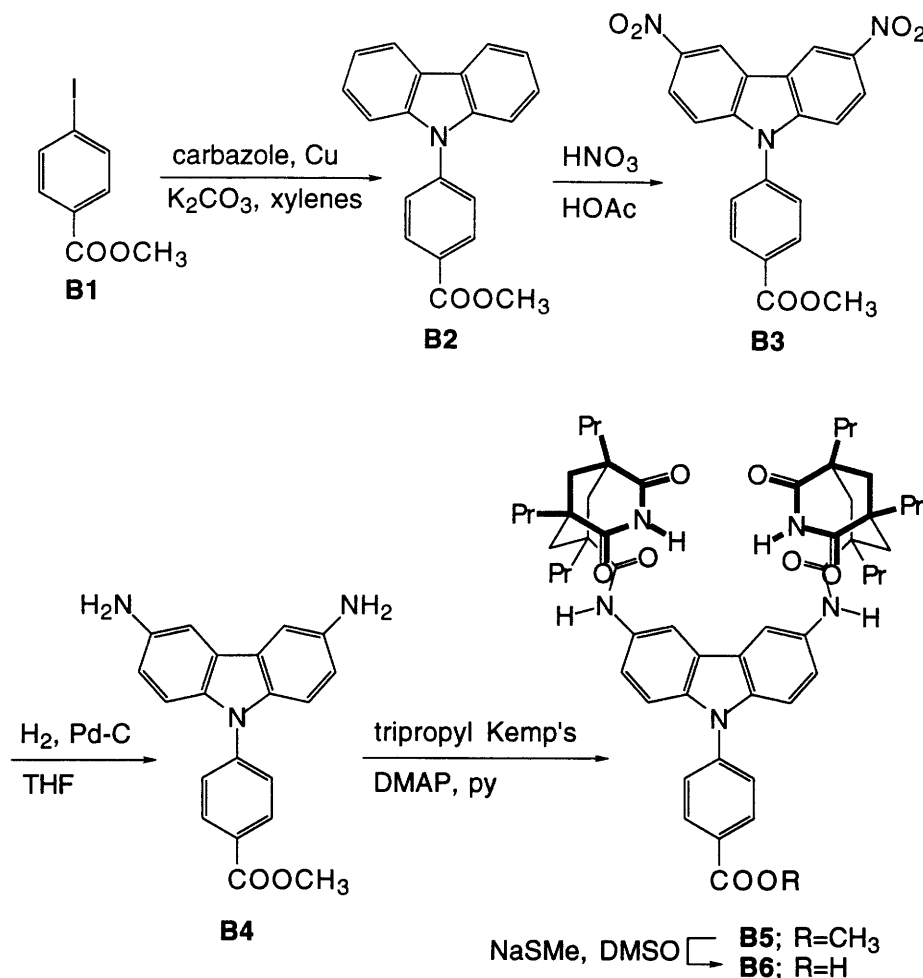


Figure B.2.2. Predicted structure of **B9** complexed with 3',5'-cAMP (stereoview).

Figures B.2.1 and B.2.2 show the predicted structures of the complex between both cyclic adenosine monophosphate derivatives and a diimide-guanidinium receptor containing a phenyl spacer. These structures were generated by minimization of structures sampled during the course of stochastic dynamics simulations in continuum aqueous solution.⁷⁰ In the complex with 2',3'-cAMP, there is a hydrogen bond between the phosphoryl groups and the guanidinium NH groups in a salt-bridge geometry whereas with 3',5'-cAMP, the planar phosphate and guanidinium functional groups are perpendicular. Only a single phosphate oxygen can bound, in a bifurcated sense, by the guanidinium in this complex. It is noteworthy that structures with bifurcated hydrogen bonding to a single oxygen also appear in the simulation of the complex of 2',3'-cAMP. That is, the salt-bridge is not the sole conformation of this complex. However, the contribution of such an interaction was predicted to be the source of selective binding of 2',3'-cAMP over 3',5'-cAMP.

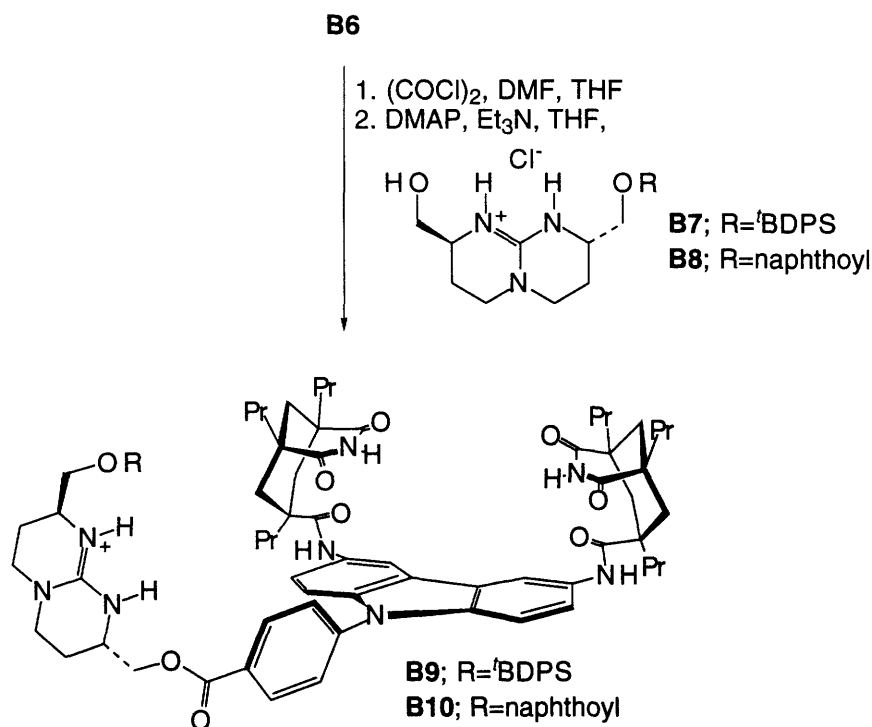


Scheme B.2.1. Synthesis of phenylscorpion.

The synthesis of a phenyl-substituted diimide cleft is summarized in Scheme B.2.1. Previous synthetic investigations by Neil Branda showed us how to place substituted phenyl rings on carbazole. Ullman coupling⁸⁶ of methyl 4-iodobenzoate with carbazole gave the substituted carbazole **B2**. Nitration of this with nitric acid proceeded easily to dinitro **B3**. Hydrogenation in previous conditions (Section A.2) were accompanied by much decomposition. Similarly, reduction with Fe/HCl in ethanol gave only moderate yield of diamine **B4**. However, hydrogenation simply in pure THF as a solvent gave a satisfactory yield of pure diamine. Acylation with the tripropyl Kemp's imide acid chloride²² gave the phenyl-substituted cleft **B5**. The free acid can be satisfactorily generated by hydrolysis of the methyl ester, but S_N2-type dealkylation⁸⁷ with methyl thiolate proceeded much more cleanly with no competing cleavage of the aromatic amide.

The cyclic AMP cleft was generated by conversion of acid **B6** to the corresponding acyl chloride with oxalyl chloride and catalytic DMF. Thionyl chloride was completely unsuitable for this purpose and led to destruction of the imide group. Acylation of the mono-protected bicyclic guanidinium⁸⁸ with the acid chloride gave the cyclic adenosine monophosphate receptor **B9**. The silyl-protected receptor **B9** was used for all titration studies. The stability of **B9** in methanol is significantly enhanced relative to receptor **I** – no degradation of the ester occurred

over the course of hours. Receptor **B10** was assembled using a naphthoylated bicyclic guanidinium diol as an analog of receptor **I** for use in transport studies (Section B.4).



Scheme B.2.2. Assembly of guan-phenylscorp.

The binding affinity of **B9** for 9-ethyladenine and the cyclic-AMP isomers in CD₃OD (Table B.2.1) was determined using the 'reverse' ¹H NMR titration protocol of Section A.2 by adding the receptor to a solution of the nucleotide. The binding affinity of the diimide cleft for 9-ethyladenine is -3.25 kcal/mol. This is the sum of hydrogen-bonding and aromatic stacking contributions in methanol. The cAMP isomers have affinity 1.1-1.6 kcal/mol higher than this as a result of the attractive phosphate-guanidinium electrostatic interaction. The predicted selectivity of the receptor manifests itself as a 0.5 kcal/mol preference for binding 2',3'-cAMP over 3',5'-cAMP.

	nucleotide	K _a (M ⁻¹)	ΔG (kcal/mol)	ΔΔG (kcal/mol)
¹ H method	2',3'-cAMP	3825	-4.88	
	3',5'-cAMP	1557	-4.35	0.53
	9-ethyladenine	241	-3.25	
³¹ P method	2',3'-cAMP	396	-3.54	
	3',5'-cAMP	241	-3.25	0.29

Table B.2.1. Association constants of receptor **B9** with cyclic AMP (Na⁺ salt) in CD₃OD.

As the selectivity is based upon two separate titrations with no control of pH or ionic strength in the ¹H NMR titrations, a competition titration was performed in CD₃OD by simultaneously monitoring the ³¹P NMR signals of *both* cAMP isomers at 25 mM ionic strength (NaBr). The monophosphate isomers have dramatically

different chemical shifts in this solvent system -- 143.093 ppm (2',3'-cAMP) and 122.222 ppm (3',5'-cAMP) as referenced to PPh₃ (118 ppm). This titration was performed by adding the receptor to the solution containing both cyclic-AMP isomers (maintained at 1 mM each) and monitoring the ³¹P chemical shift of the phosphate group. Only small (0.3 and 1.0 ppm) chemical shift changes occur over the course of the titration because phosphate oxygens are already significantly hydrogen-bound in methanolic solution. On complexation, only a small net change in hydrogen-bonding occurs. Despite these small changes, the titration curve of each phosphate fits the 1:1 binding isotherm (Equation A.2.11). The strength of association is considerably reduced (>1 kcal/mol) as a result of the high ionic strength, which reduces the influence of the ionic interaction on binding relative to the hydrogen-bonding component. Nevertheless, as seen in the ¹H NMR titration, there is again a preference for binding to 2',3'-cAMP, but of only 0.29 kcal/mol.

It is interesting that the affinity of **B9** for 3',5'-cAMP in 25 mM salt is the same as for 9-ethyladenine without control of the ionic strength. However, since hydrogen-bonding does have an electrostatic component,⁴⁵ it would be unwise to over-interpret this to mean that no charge interaction remains in the complex with 3',5'-cAMP at 25 mM ionic strength.

Titration of the cAMP isomers with a benzoylated bicyclic guanidinium gave association constants were less than 10 M⁻¹. The chemical shift changes with the free guanidinium were much smaller than those when the cleft is present and fit to a 1:1 binding isotherm is very poor. This indicates that the effects observed in the ³¹P titration arise from binding within the cleft and not as a result of changes in the medium from the addition of guanidinium.

To fully ascertain that the positioning of the bicyclic guanidinium is responsible for the observed selectivity, the affinity of both cAMP isomers to methyl ester **B5** should be known. It is possible that the phenylcarbazole-diimide receptor itself has an inherent *shape* selectivity for 2',3'-cAMP. Unfortunately, methyl ester **B5** is not sufficiently soluble in methanol for NMR titration. Transfer of the system to a similar polar solvent, acetone, in order to measure all of these interactions was not possible due to insolubility of the cyclic nucleotides.

In conclusion, we have been able to construct a receptor with specificity for 2',3'-cAMP on the basis of a prediction from molecular modeling. The source of the specificity is predicted to arise from hydrogen-bonding within a salt-bridge interaction, but could not be experimentally verified.

This receptor system can be transferred to aqueous systems using the tris((benzyloxy)methyl) derivative of Kemp's triacid.⁴² Yoko Kato has made a water-soluble versions of **B5** and **B9** and studied their affinity for adenosine and cyclic adenosine monophosphates in water.⁴³ The selectivity of 2',3'-cAMP is reduced to 0.2 kcal/mol in this solvent system (at 50 mM ionic strength).

The strength (enthalpy) of a phosphate-guanidinium interaction in water can be extracted from the binding data by making use of a free energy pathway to subtract all of the contributions from hydrogen-bonding and hydrophobic interactions. Isolated guanidinium and phosphates show extremely low affinity (-0.2 kcal/mol⁸⁹)

for each other in free solution. This value is free of contributions from intermolecular translational and rotational entropy as it based upon comparisons within a complex. The values obtained from this kind of study are more representative of interactions in free aqueous solution than those determined from protein mutagenesis studies²⁴ where the specialized, highly specific environment of a protein interior can be significantly different from aqueous solution.

3. TRINUCLEOTIDE RECOGNITION

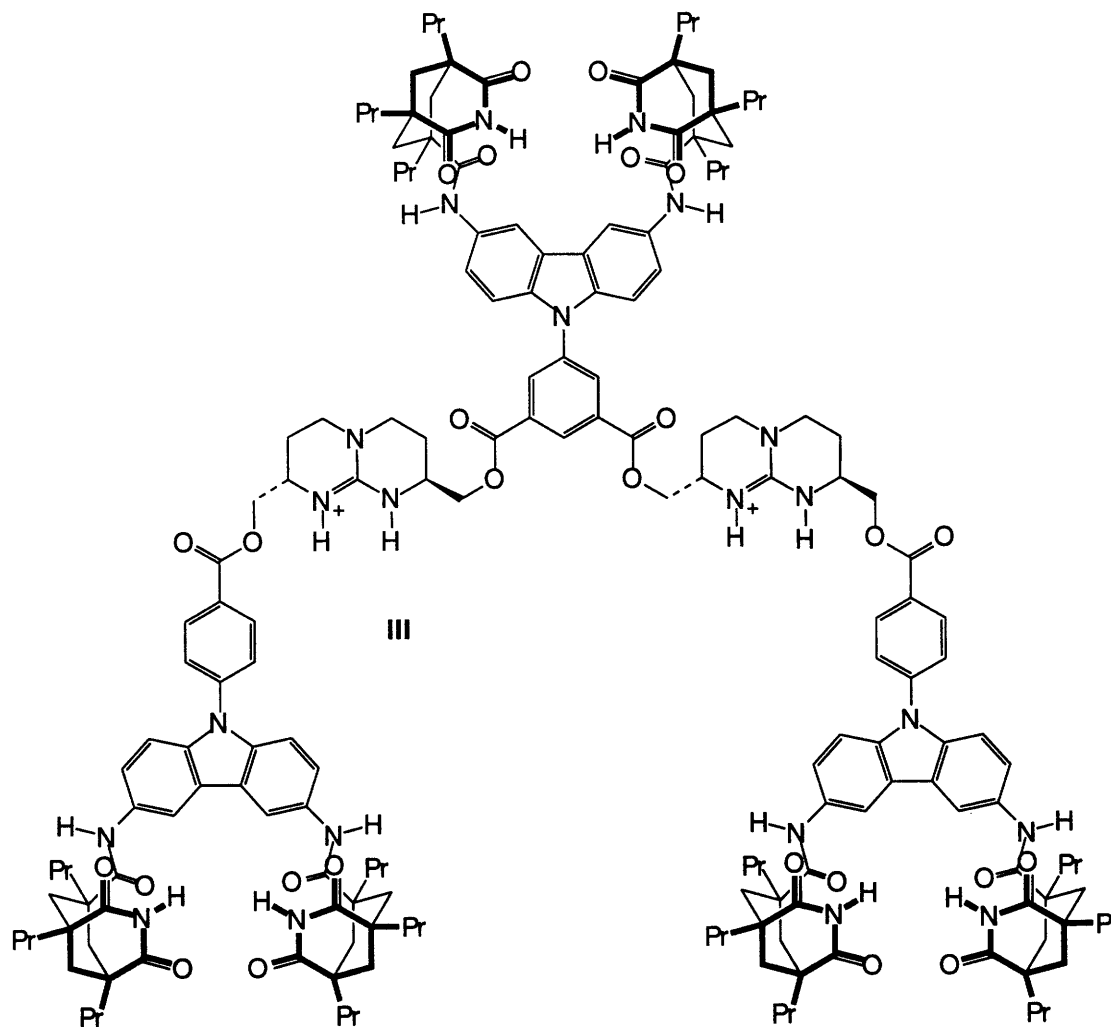
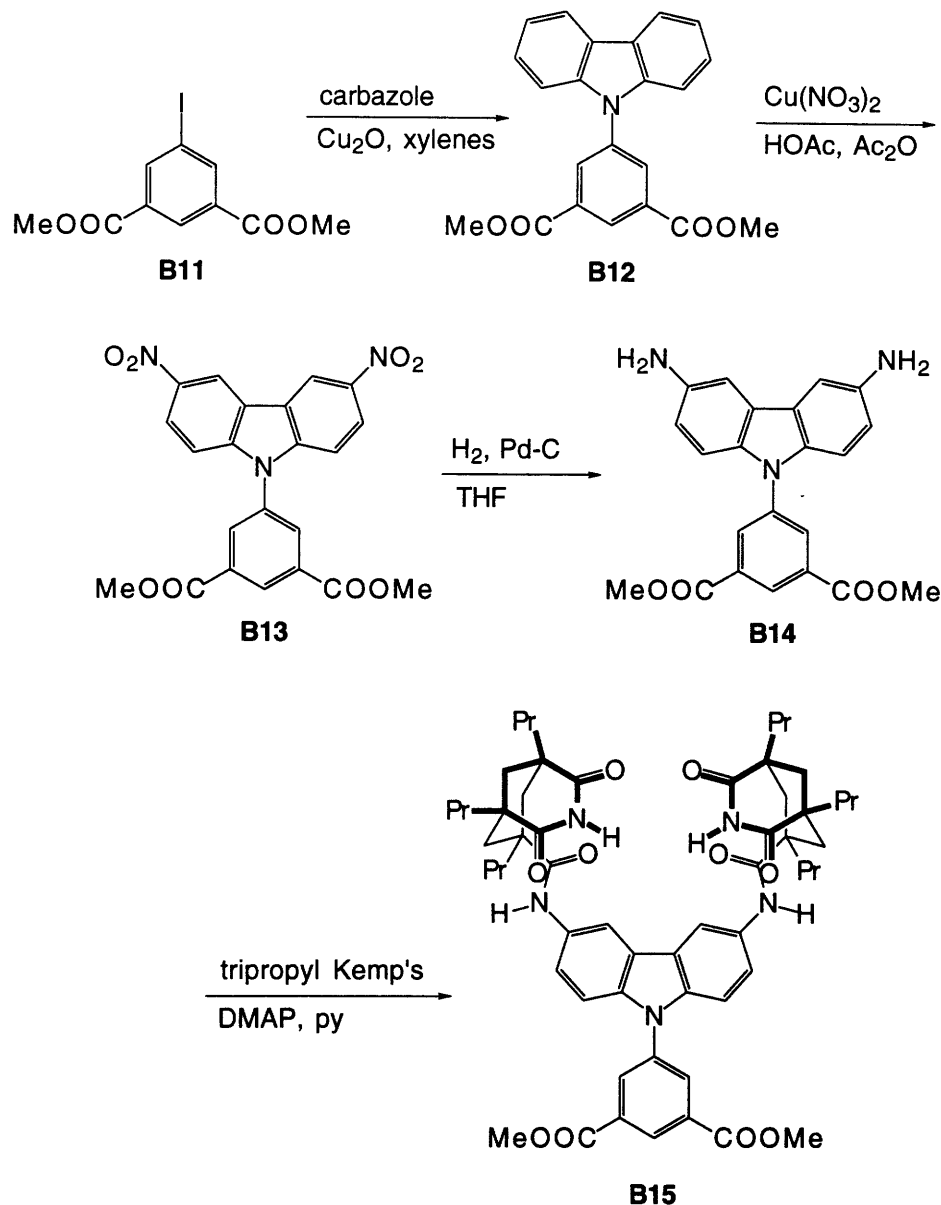


Figure B.3.1. An adenosine trinucleotide receptor.

The utility of phenyl-substituted carbazole-diimide receptors for nucleotide recognition can be extended by exploiting the rigid shape to provide divergence. For the recognition of longer nucleotides (*i.e.* trinucleotides), two phosphate-binding guanidinium units and three adenosine recognition elements must be linked. A very simple way to do this is to have a central adenine-binding cleft which joins two adenosine mononucleotide-binding clefts of the form of **B9**. In collaboration with Javier de Mendoza, receptor **B15** was proposed for this purpose. The isophthaloyl substituent on carbazole N⁹ provides a method for linking two **B9**-like receptors

together to provide recognition sites for two phosphates and three adenines, as shown in Figure B.3.1. The keystone of receptor **III** is the central isophthaloyl-substituted diimide cleft. The two flanking units have already been discussed in Section B.2.



Scheme B.3.1. Synthesis of isophthaloylscorpion.

The synthesis of **B15** is considerably more difficult than any of the other aromatic-substituted carbazole-diimide receptors discussed here (Sections B.2 and C.3). Dimethyl 5-iodoisophthalate and carbazole were coupled using Ullman-type conditions⁸⁶ with copper(I) oxide. The reaction proceeded not at all using copper powder and potassium carbonate. It is a general observation that cuprous oxide is a more successful reagent for Ullman-type coupling of *meta*-substituted aromatic rings while copper powder is more appropriate for *para*-substituted aromatic rings. The isophthaloyl carbazole **B12** behaved much differently than previous N^9 -

aromatic-substituted carbazole molecules. Nitration of this was very difficult, and complicated by extremely poor solubility. Reaction with nitric acid did not cleanly give the 3,6-dinitro product. Harsher conditions led to hydrolysis of the esters and tri-nitration of the carbazole. Finally, partially pure dinitrocarbazole **B13** could be produced by nitration with cupric nitrate. Purification was not possible.

Reduction of **B13** was a major stumbling block. Any reduction conditions with alcohol present (Section A.2 or Fe/HCl) gave no appreciable yield of **B15** after the acylation reaction. However, the use of THF as solvent allowed hydrogenation to occur cleanly. Diamine **B14** is even more unstable than diamines **A3**, **B4**, or **C12** and coupling of this with tripropyl Kemp's imide acid chloride²² gives a lower yield of **B15** than for the analogous reactions. The binding affinity of this cleft with adenosine in CDCl₃ was estimated as 10⁵ M⁻¹ based on the 'reverse' titration protocol of Section A.2.

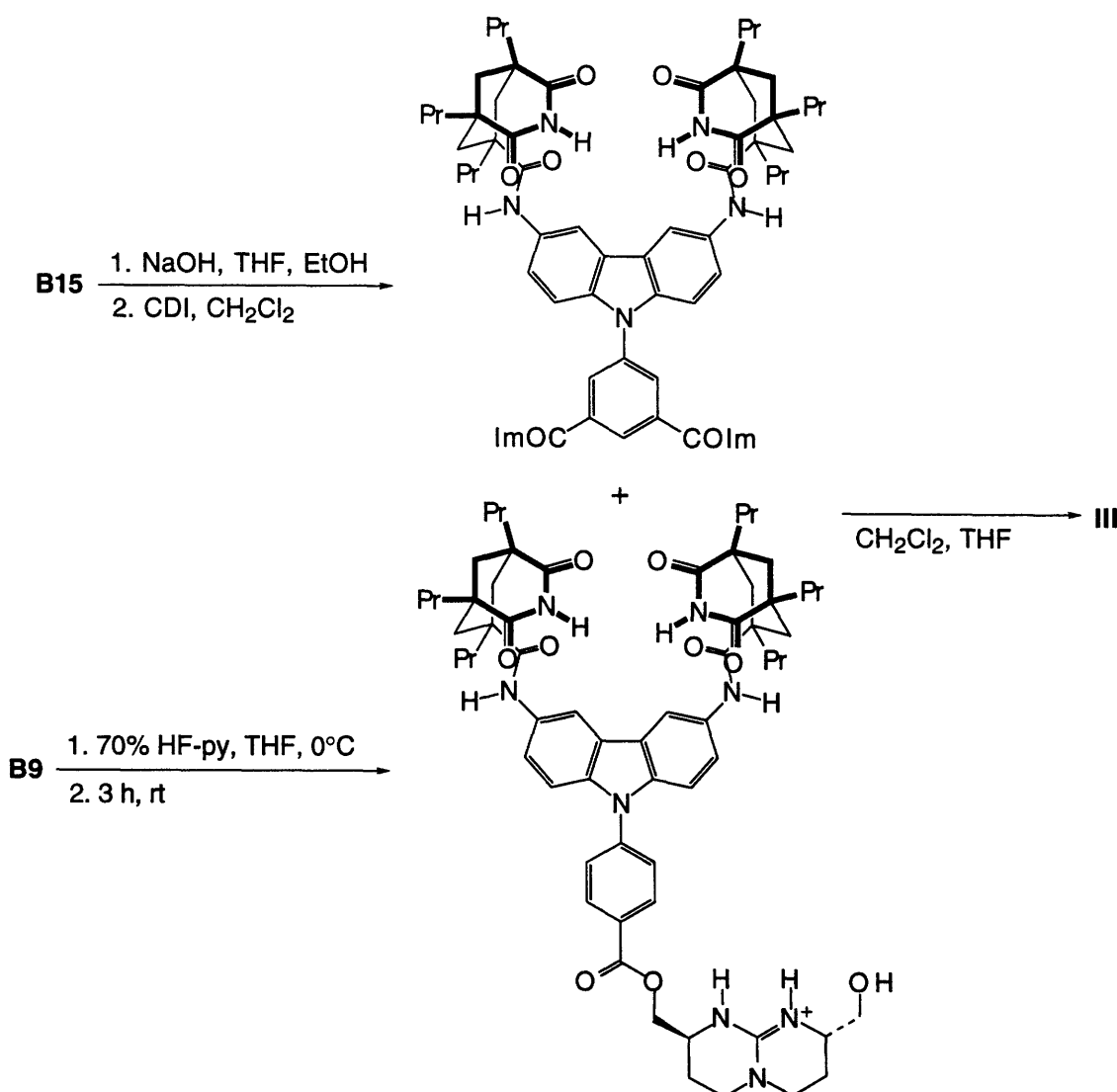


Figure B.3.2. Assembly of trinucleotide receptor **III**.

The trinucleotide receptor **III** was assembled at the Universidad Autonoma by Armando Salmerón in Javier de

Mendoza's laboratory by hydrolysis of diester **B15** and activation with carbonyl diimidazole, followed by acylation of deprotected **B9** to give receptor **III** (Figure B.3.2).

Receptor **III** is an extremely large molecule with multiple possible conformations. With 15 free rotors in the receptor and 9 free rotors in its substrate, ApApA, the conformational flexibility of the complex makes modeling and structural studies very difficult. Currently, ^1H NMR investigation into the structure of the complex between **III** and ApApA is in progress at the Universidad Autonoma (600 MHz ROESY).

4. NUCLEOTIDE TRANSPORT

A number of studies have been performed on the transport of nucleotides and nucleotide derivatives across liquid membranes. The receptors used for this purpose have included quaternary ammonium salts^{80,81} and expanded porphyrin derivatives.^{82,90} Early imide receptors from this laboratory have been used for the transport of nucleoside derivatives across liquid membranes.¹⁶ With the addition of phosphate recognition, it is now possible to transport *nucleotides* through liquid membranes using imide-based adenine receptors.

The transport of nucleotides across a liquid membrane is studied using a U-tube apparatus, shown in Figure B.4.1. All transport studies were performed by Cecilia Andreu.⁹¹ The source phase contains an aqueous solution of the nucleotide and the receiving phase is an aqueous salt solution, while the organic phase is dichloroethane. As the amount of transport is affected by mixing at the organic-aqueous interface, stirring is maintained at a constant rate.

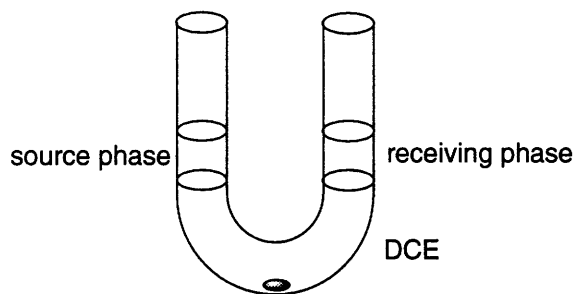


Figure B.4.1. U-tube apparatus for liquid membrane transport studies.

The transport that occurs through the liquid membrane is referred to as facilitated transport.⁹² That is, simple diffusion is not possible; nucleotides will *not* move through the bulk organic phase without the assistance of a carrier. The concentrations of nucleotide in the source and receiving phase are assayed by UV/Vis spectroscopy and typically follow the curves shown in Figure B.4.2. The concentration in the source phase decreases as that in the receiving phase increases by the same amount. There are two concentration gradients -- a nucleotide ion gradient from the source to receiving phase and a chloride ion gradient from the receiving to the source phase. The positively-charged carrier must exist as a salt in organic solution and so acts as a transport agent for both of these anions. At the source-phase interface, the carrier releases chloride and complexes with nucleotide; the converse occurs at the interface with the receiving phase. Initially, then, there are two driving forces for

nucleotide transport. However, once the nucleotide concentration reaches equilibrium between the source and receiving phase, it is the continuing movement of chloride ions that drives nucleotide transport *against* its concentration gradient. This is active transport. The rates of initial and active transport are approximated from the most linear data points immediately after time zero and equivalence, respectively.

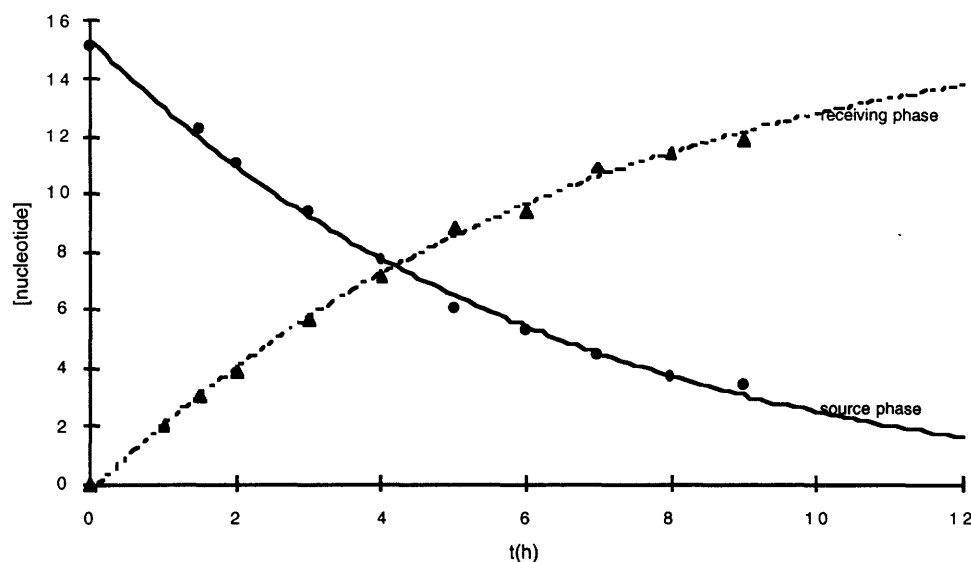


Figure B.4.2. Liquid membrane transport concentration profile.

4.1 Mononucleotides

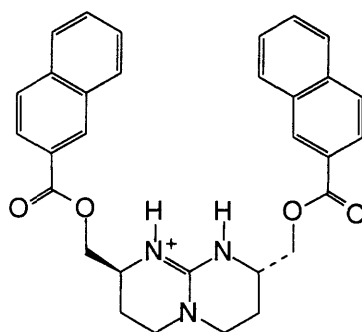
The ability of receptor **I** to transport nucleotides through a dichloroethane liquid membrane was examined, using a 10 mM NaCl solution in the receiving phase. The results of the transport study (Table B.4.1) show that receptor **I** transports cAMP derivatives preferentially over the corresponding guanosine nucleotides. No transport of adenosine monophosphate nucleotides is seen when adenine recognition is removed as in receptor **IV** (Figure B.4.3). The lack of transport of the guanosine nucleotides reflects the lower affinity of the diimide clefs for guanine (Section A.2.3) coupled with the much lower organic solubility of guanine.

[carrier] (mM)	nucleotide (15 μ M)	initial transport (pmol/h-cm ²)	active transport (pmol/h-cm ²)
7.5	2',3'-cAMP	2.194 \pm 0.016	1.216 \pm 0.004
	3',5'-cAMP	1.820 \pm 0.040	0.964 \pm 0.013
	2',3'-cGMP	0	0
	3',5'-cGMP	0	0
25	3'-AMP	1.383 \pm 0.037	0.685 \pm 0.023
	5'-AMP	0	0

Table B.4.1. Transport of mononucleotides by carrier **I**.

There is some selectivity in the transport of adenosine monophosphates. The positioning of phosphate on the

5'-end reduces the effectiveness of transport. That is, 5'-AMP is not at all transported and 3',5'-cAMP is transported less efficiently than 2',3'-cAMP.



IV

Figure B.4.3. Transport control.

Receptor **B10** is an analog of receptor **I** with a phenyl spacer in place of methylene. This could have multiple effects including selectivity as a result of reduced flexibility and improved positioning and additional solubilization of hydrophilic nucleotides as a result of greater hydrophobic contacts with the phenyl ring.

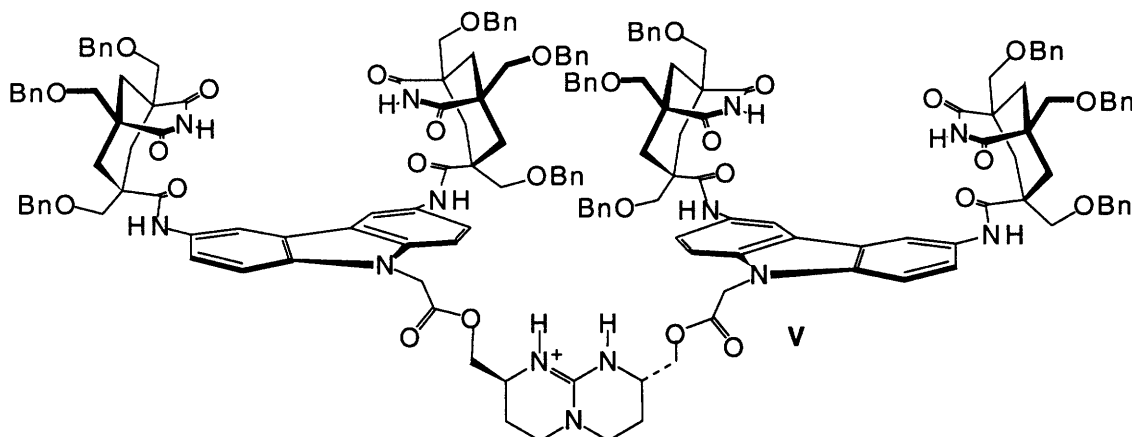
Transport studies with **B10** support these expectations (Table B.4.2). The receptor is more effective at transporting nucleotides than receptor **I**. Under identical conditions (Table B.4.2, line 5), **B10** is 55% more efficient in the initial transport of 3'-AMP and 2.5 times more efficient in active transport. The binding selectivity of receptor **B9** for 2',3'-cAMP manifests itself as a *decrease* in transport (Table B.4.2, lines 2 and 3). Significantly higher salt concentration is required in the receiving phase in order to extract bound 2',3'-cAMP from the organic layer. That is, the receptor can efficiently take up 2',3'-cAMP from the source phase, but because of its binding affinity, release to the receiving phase is more difficult. In addition, this receptor is now capable of transporting cGMP derivatives (Table B.4.2, lines 6 and 7). As no transport of these was seen at all for receptor **I**, neither the guanidinium nor the diimide, but increased hydrophobic contacts with the phenyl group, are responsible. The greater hydrophobic nature of carrier **B10** allows the possibility for transport of a greater variety of nucleotides across liquid membranes.

[carrier] (mM)	nucleotide (15 μ M)	receiving phase [NaCl] (mM)	initial transport (pmol/h-cm ²)	active transport (pmol/h-cm ²)
4.5	2',3'-cAMP	10	1.337 \pm 0.012	0.536 \pm 0.13
7.5	"	250	2.438 \pm 0.036	1.419 \pm 0.015
	"	500	2.334 \pm 0.063	1.476 \pm 0.068
	3',5'-cAMP	10	2.203 \pm 0.093	1.518 \pm 0.001
25	3'-AMP	10	2.151 \pm 0.077	1.699 \pm 0.004
35	2',3'-cGMP	10	2.246 \pm 0.017	1.387 \pm 0.039
	3'5'-cGMP	10	2.442 \pm 0.029	1.448 \pm 0.122

Table B.4.2. Transport of mononucleotides by **B10**.

4.2 Dinucleotides

Having demonstrated the capability of linked diimide-guanidinium receptors to act as carriers for mononucleotides, the transport of dinucleotides was examined. For this purpose, receptor **V** was used. The results of these studies are summarized in Table B.4.3.

Figure B.4.4. Dinucleotide carrier **V**.

With a salt concentration in the receiving phase of 10 mM, receptor **V** was able to transport adenosine-containing dinucleotides with varying efficiency. The deoxyadenosine dinucleotide was transported somewhat slower than the corresponding mononucleotide (Table B.4.3, line 1). Transport of cytidine-containing dinucleotides or the guanosine dinucleotide was not possible (Table B.4.3, lines 3,4, 5, and 9) but guanosine could be transported as the mixed dinucleotide (Table B.4.3, line 8).

The improved transport ability of receptors bearing a phenyl group on carbazole N⁹ will be exploited for dinucleotide transport in a similar fashion to the mononucleotide system. These studies will be carried out by Yoko Kato and Cecilia Andreu. It would be expected that a wider variety of dinucleotides will be carried through the liquid organic membrane by a receptor of this type.

[carrier] (mM)	nucleotide (10 μ M)	initial transport (μ mol/h-cm ²)	active transport (μ mol/h-cm ²)
7.5	d(AA)	1.027 \pm 0.004	0.755 \pm 0.018
10	d(AC)	0	0
	d(CA)	0	0
	d(CG)	0	0
12	dApT	1.058 \pm 0.053	0.792 \pm 0.041
	TpdA	0.910 \pm 0.001	0.484 \pm 0.002
20	d(AG)	0.360 \pm 0.004	-
	d(GG)	0	0

Table B.4.3. Transport of dinucleotides by carrier **V**.

4.3 Trinucleotides

Keeping in mind the lessons of adenosine recognition and hydrophobic contact from smaller systems, we examined the transport of trinucleotides with receptor **III**. The two bicyclic guanidinium units serve to neutralize the charge of the two phosphates of a trinucleotide, and the adenosine recognition elements convey hydrogen-bond specificity to adenine. As in the simpler systems (Sections B.4.1 and B.4.2), the extended aromatic systems of receptor **III** should also contribute to the transport of non-adenosine nucleotides.

carrier	[carrier] (mM)	receiving phase [NaCl] (mM)	initial transport (pmol/h-cm ²)	active transport (pmol/h-cm ²)
III	5	250	0.621±0.041	0.556±0.029
		500	1.018±0.006	0.834±0.021
VI	20	10	1.075±0.008	0.781±0.038

Table B.4.4. Transport of ApApA (10 μM).

Preliminary results from studies of the transport of the adenine trinucleotide through liquid membranes show that **III** is only a poor carrier of ApApA. The receptor, it seems, is too good at its job. Uptake into the organic phase is very efficient, but release into the receiving phase requires high salt concentration. The interaction between receptor **III** and ApApA is so strong in organic solvent that the nucleotide remains in that phase.

Efficient transport requires a compromise in binding affinity such that release occurs after the initial extraction.

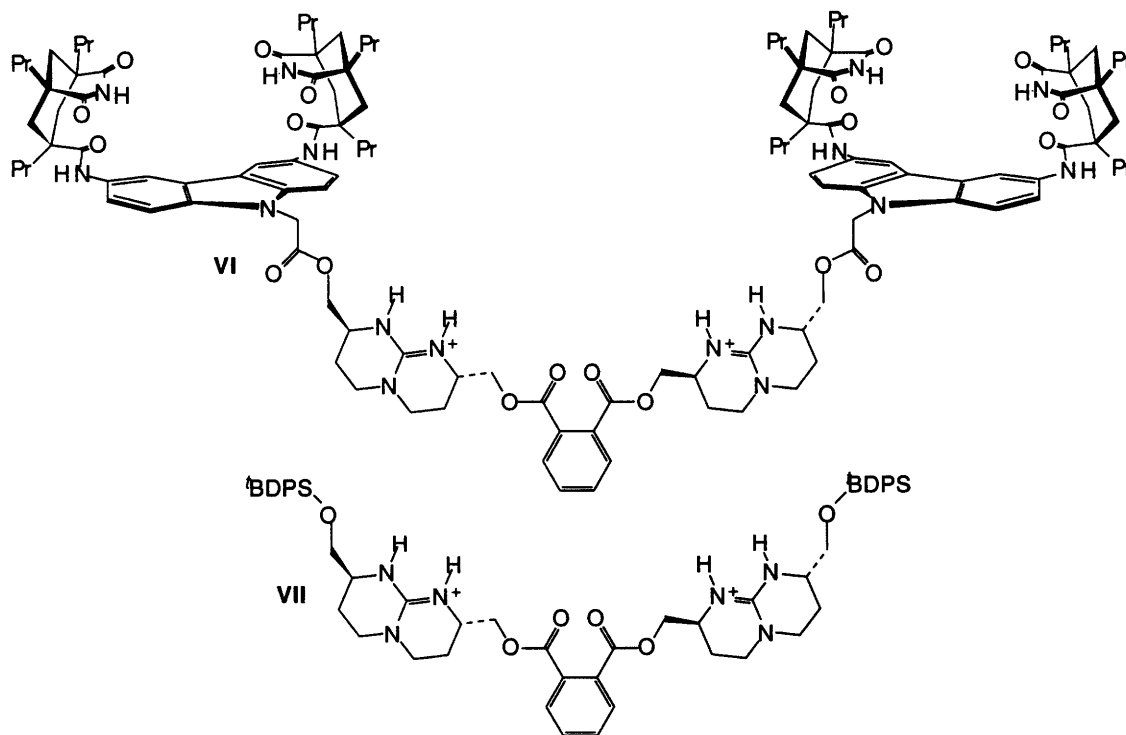


Figure B.4.5. Controls for trinucleotide transport.

Receptor **VI** was prepared in Javier de Mendoza's laboratory and found to be much more effective at the transport

of ApApA through a dichloroethane liquid membrane. With only methylene spacer elements and lacking one adenine-binding unit, the binding affinity for ApApA is moderated so that transport is possible. Charge-stabilization alone, however, is not sufficient for transport as control **VII**, lacking all adenine-binding capability, is ineffective for transport.

The overall lessons of these transport studies are that, in general, high sequence specificity is not desired as this reduces the range of application for a carrier. Furthermore, the strength of molecular recognition in the organic phase must be moderated such that transport, rather than mere extraction, occurs. Proceeding from this, investigation into the transport of nucleotides of varying sequence and length are in progress.

5. EXPERIMENTAL

5.1 Synthesis

General

All commercially available compounds (Aldrich) were used without further purification unless otherwise indicated. DMSO- d_6 (99.9% D, MSD Isotopes or Cambridge Isotope Labs) were used from freshly opened vials. $CDCl_3$ (99.8% D, MSD or CIL) was stored over molecular sieves and passed through dry basic alumina just prior to use. Tetrahydrofuran used in anhydrous conditions were distilled from sodium-benzophenone ketyl; dimethyl sulfoxide was distilled from sodium hydroxide and stored over molecular sieves (4 Å). Thin layer chromatography was performed on Merck Silica 60 F-254 precoated TLC plates. Column chromatography was performed on Merck Silica Gel 60 (230-400 mesh) according to Still *et al.*⁵⁹ Glassware used for anhydrous conditions was either 1) baked overnight at 150 °C, assembled hot, cooled under vacuum, and filled with argon; or 2) was flamed dry under vacuum, cooled and filled with argon before use. Standard inert-atmosphere techniques were used for solvent transfers by syringe.

1H NMR spectra were recorded on Varian XL-300, UN-300, and VXR-500 spectrometers in solvents as indicated. Chemical shifts are reported as parts per million (δ) relative to residual solvent protons. Melting points were obtained on an Electrothermal IA9100 digital melting point apparatus and are calibrated. IR spectra were recorded on a Perkin-Elmer series 1600 FT-IR spectrometer. High resolution mass spectra were obtained on a Finnegan MAT 8200 mass spectrometer.

Methyl 4-Iodobenzoate (B1)

Thionyl chloride (10 mL) was added to a suspension of 4-iodobenzoic acid (7.54 g, 30.4 mmol) in dry MeOH over 10 min. The reaction was refluxed for 1 h under a drying tube, cooled to room temperature and stirred for 22 h. Concentration and drying under vacuum afforded quantitative yield of methyl ester **B1** which was used without further purification: 1H NMR (300 MHz, DMSO- d_6) δ 7.906 (d, 2 H, J = 8.7 Hz), 7.691 (d, 2 H, J = 8.4

Hz), 3.832 (s, 3 H).

N-(4-(Methoxycarbonyl)phenyl)carbazole (**B2**)

Carbazole (7.90 g, 30.15 mmol) and methyl 4-iodobenzoate **B1** (1.016 equiv) were intimately mixed in a mortar and pestle with K₂CO₃ (1.286 equiv) and catalytic Cu powder (202 mg). The mixture was refluxed with xylenes (10 mL) with magnetic stirring for 2 days. The reaction was stopped and benzene (250 mL) was added while the mixture was still warm (and liquid). The suspension was refluxed for 3 h, cooled, and filtered through Celite. Solvent was removed by rotary evaporation and the resulting brown oil left to crystallize overnight. The crude product was recrystallized from refluxing MeOH (200 mL) and EtOAc (10 mL) to give 6.057 g of white crystals. The mother liquor was concentrated and recrystallized from MeOH to give 1.79 g of white crystals. The combined yield of **B2** was 7.847 g (86%): mp 120.7-121.5 °C; IR (KBr) 3050,2946, 1721, 1602, 1513, 1452, 1432, 1290, 1280, 1225, 1168, 1115, 1102, 767, 736, 723, 700 cm⁻¹; ¹H NMR (500 MHz, CDCl₃) δ 8.289 (d, 2 H, J = 8.5 Hz), 8.149 (dd, 2 H, J = 7.5, 0.5 Hz), 7.689 (d, 2 H, J = 8.0 Hz), 7.477 (dd, 2 H, J = 8.2, 0.7 Hz), 7.430 (dt, 2 H, J = 7.6, 1.0 Hz), 7.319 (dt, 2 H, J = 7.2, 0.8 Hz), 3.995 (s, 3 H); HRMS (EI) calcd for C₂₀H₁₅NO₂, 301.1103; found, 301.1100.

N-(4-(Methoxycarbonyl)phenyl)-3,6-dinitrocarbazole (**B3**)

Concentrated nitric acid (78 mL) was added dropwise to a suspension of carbazole **B2** (7.63 g, 25.32 mmol) in HOAc (120 mL) at 50 °C. Additional HOAc (50 mL) was added to thin the suspension and the temperature raised to 70 °C. The suspension was stirred at this temperature for 11 h and cooled to room temperature for 2 h. Product was precipitated by addition of water (100 mL), isolated by filtration, and dried in a vacuum desiccator to give 8.819 g (89%) of pale yellow powder **B3**, which was used without further purification: ¹H NMR (300 MHz, DMSO-d₆) δ 9.654 (d, 2 H, J = 2.1 Hz), 8.421 (dd, 2 H, J = 9.1, 2.2 Hz), 8.297 (d, 2 H, J = 8.4 Hz), 7.913 (d, 2 H, J = 8.4 Hz), 7.621 (d, 2 H, J = 9.3 Hz), 3.941 (s, 3 H).

N-(4-(Methoxycarbonyl)phenyl)-3,6-diaminocarbazole (**B4**)

Dinitro carbazole **B3** (224 mg, 0.5724 mmol) and 10% Pd-C (93 mg, 42 wt%) were suspended in THF (25 mL) and hydrogenated under balloon pressure for 14 h. The suspension was filtered through Celite. An unstable yellow solid was recovered by rotary evaporation, dried under vacuum, and used without further purification: ¹H NMR (250 MHz, CDCl₃) δ 8.223 (d, 2 H, J = 8.6 Hz), 7.632 (d, 2 H, J = 8.6 Hz), 7.326 (d, 2 H, J = 2.7 Hz), 7.315 (d, 2 H, J = 8.2 Hz), 6.820 (dd, 2 H, J = 8.7, 2.1 Hz), 3.970 (s, 3 H).

N-(4-(Methoxycarbonyl)phenyl)-3,6-bis(((cis,cis-2,4-dioxo-1,5,7-tripropyl-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino)carbazole (**B5**)

The above diamine **B4** (0.5724 mmol, theoretical) was refluxed with tripropyl Kemp's imide acyl chloride²² (1.17 mol equiv) in pyridine (10 mL) with catalytic DMAP (0.04 equiv) for 7 h under an Ar balloon. Solvent was

removed by rotary evaporation under reduced pressure. The residue was taken up in EtOAc (30 mL) and washed with 1% HCl (10 mL), water (10 mL), and brine (10 mL) and dried over Na₂SO₄. After *gravity* chromatography (10-20% EtOAc/CH₂Cl₂), the cleanest fractions were triturated in MeOH (~3 mL) to afford 373 mg (69%) of pale beige powder after filtration and drying: IR (KBr) 3379, 3214, 2958, 2871, 1696, 1605, 1514, 1466, 1281, 1196, 1105 cm⁻¹; ¹H NMR (300 MHz, DMSO-d₆) δ 10.366 (s, 2 H, imide), 9.247 (s, 2 H, amide), 8.302 (d, 2 H, J = 1.5 Hz), 8.230 (d, 2 H, J = 8.1 Hz), 7.790 (d, 2 H, J = 8.1 Hz), 7.503 (dd, 2 H, J = 9.0, 1.8 Hz), 7.414 (d, 2 H, J = 9.0 Hz), 3.915 (s, 3 H), 2.678 (d, 4 H, J = 13.5 Hz), 2.029 (d, 2 H, J = 12 Hz), 1.78 (m, 4 H), 1.51 (m, 4 H), 1.4-1 (m, 22 H), 0.874 (t, 12 H, J = 6.9 Hz), 0.802 (t, 6 H, J = 7.3 Hz); HRMS (EI) calcd for C₅₆H₇₁N₅O₈, 941.5303; found, 941.5304.

N-(4-Carboxyphenyl)-3,6-bis[*[(cis,cis-2,4-dioxo-1,5,7-tripropyl-3-azabicyclo[3.3.1]non-7-yl)carbonyl]amino*]carbazole (**B6**)

Ester **B5** (174 mg, 0.1847 mmol) and methyl thiolate (81 mg, 6.26 equiv) were dissolved in anhydrous DMSO (5 mL) under argon at 78 °C in an oil bath. The reaction was stirred at this temperature for 1 h, cooled, and quenched with 1 N HCl (10 mL). The suspension was extracted with CHCl₃ (50 mL) and the organic layer was washed with 1 N HCl (2 x 20 mL), water (30 mL), and brine (30 mL), and dried over Na₂SO₄. Solvent was removed by rotary evaporation and the off-white solid dried under vacuum to give 168 mg (98%) of powder which was used without further purification: ¹H NMR (300 MHz, DMSO-d₆) δ 10.341 (s, 2 H, imide), 9.232 (s, 2 H, amide), 8.299 (d, 2 H, J = 1.6 Hz), 8.206 (d, 2 H, J = 8.6 Hz), 7.747 (d, 2 H, J = 8.5 Hz), 7.499 (dd, 2 H, J = 9.0, 1.6 Hz), 7.403 (d, 2 H, J = 8.8 Hz), 2.681 (dd, 4 H, J = 14.7, 0.9 Hz), 2.029 (d, 2 H, J = 12.9 Hz), 1.78 (m, 4 H), 1.51 (m, 4 H), 1.4-1 (m, 22 H), 0.876 (t, 12 H, J = 6.9 Hz), 0.805 (t, 6 H, J = 7.2 Hz).

(2*S*,8*S*)-2-Hydroxymethyl-8-[[*(tert-butyl)diphenylsilyloxy*]methyl]-3,4,6,7,8,9-hexahydro-2*H*-pyrimido[1,2-*a*]pyrimidine Hydrochloride (**B7**)

The ¹BDMS, ¹BDPS protected diol⁸⁸ (2000 mg, 0.3399 mmol) was dissolved in THF (3 mL), water (3 mL), and HOAc (10 mL) and stirred for 23 h at room temperature. Solvent was removed by rotary evaporation at room temperature. Three times, toluene (~5 mL) was added and rotavapped to remove acetic acid. A white powder was formed by slow evaporation of a hexanes/CH₂Cl₂ solution. The solid was triturated in hexanes (~3 mL) and the supernatant removed after settling. This was repeated and the solid dried under vacuum to give quantitative yield of alcohol **B7**, which was used without further purification: ¹H NMR (500 MHz, CDCl₃) δ 9.185 (s, 1 H, NH), 7.696 (s, 1 H, NH), 7.63 (m, 4 H), 7.41 (m, 6 H), 3.810 (dd, 1 H, J = 12.0, 3.5 Hz), 3.738 (d, 1 H, J = 5.0 Hz), 3.59 (m, 3 H), 3.51 (m, 1 H), 3.32 (m, 1 H), 3.24 (m, 3 H), 2.07 (m, 1 H), 1.91 (m, 3 H), 1.070 (s, 9 H).

(2S,8S)-2-[[[2-Naphthoxy]carbonyl]methyl]-8-[[[tert-butyl(diphenylsilyloxy)methyl]-3,4,6,7,8,9-hexahydro-2H-pyrimido[1,2-a]pyrimidine Hydrochloride (**B8**)

Alcohol **B8** (48 mg, 0.1012 mmol) was dissolved in anhydrous THF (4 mL) under argon and naphthoyl chloride (1.45 equiv) was added, followed by TEA (50 μ L, 3.54 equiv). A precipitate of TEA•HCl forms within min. The suspension was stirred at room temperature for 23 h and quenched with butylamine (1 drop). After filtration, the solvent was removed by rotary evaporation and the residue purified by chromatography (5% MeOH/CHCl₃) to give 55 mg (86%) of a colorless oil that formed a powder on slow evaporation of a hexanes/CH₂Cl₂ solution, which was used without further purification.

Silyl guan-phenylscorpion (**B9**)

Acid **B6** (104 mg, 0.1120 mmol) was dissolved in anhydrous THF (4 mL) with catalytic DMF (~2 μ L) and oxalyl chloride (9.21 equiv) was added. The solution was stirred under argon for 35 min and reduced by rotary evaporation. After drying under vacuum, the acyl chloride was dissolved in anhydrous THF (10 mL) under Ar with alcohol **B7** (0.998 equiv) and TEA (19 equiv) was added. After stirring for 22 h at room temperature, the suspension was filtered and the filtrate reduced by rotary evaporation under reduced pressure. Flash chromatography (8% MeOH/CHCl₃) afforded 84 mg (56%) of an off-white powder: IR (KBr) 3265, 2958, 2871, 1701, 1628, 1465, 1271, 1198, 1112, 704 cm⁻¹; ¹H NMR (500 MHz, DMSO-d₆) δ 10.636 (s, 2 H, imide), 9.255 (s, 2 H, amide), 8.300 (m, 4 H), 7.971 (m, 1 H), 7.818 (d, 2 H, J = 8.5 Hz), 7.615 (m, 6 H), 7.52-7.39 (m, 12 H), 4.444 (m, 1 H), 4.338 (m, 1 H), 3.868 (m, 1 H), 3.649 (m, 2 H), 3.590 (m, 1 H), 2.675 (d, 4 H, J = 13.4 Hz), 2.029 (d, 2 H, J = 13 Hz), 2.06 (m, 4 H), 1.78 (m, 6 H), 1.51 (m, 4 H), 1.4-1 (m, 22 H), 1.013 (s, 9 H), 0.875 (t, 12 H, J = 6.8 Hz), 0.801 (t, 6 H, J = 7.1 Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₈₀H₁₀₃N₈O₉Si (M-Cl), 1347.7617; found, 1347.7606.

Naphthoyl guan-phenylscorpion (**B10**)

The acid chloride of acid **B6** (84 mg, 90.5 μ mol) was prepared as above. The 'BDPS-protected guanidinium **B8** (55 mg, 87.54 μ mol) was dissolved in anhydrous THF (4 mL) and treated with TBAF (1.0 M in THF, 300 μ L, 3.43 equiv) at room temperature for 2 h. The deprotected alcohol was added *in situ* via cannula to the acid chloride, followed by a THF (2 x 2 mL) rinse of the alcohol flask. After 10 min, TEA (24 μ L, 2 equiv) was added. No precipitate formed. After 5 h, solvent was removed by rotary evaporation. The residue was taken up in CHCl₃, and washed with 0.1 N HCl (15 mL), water (10 mL), and brine (10 mL), and dried over Na₂SO₄. Chromatography (8% MeOH/CHCl₃) gives 39.2 mg (34%) of pure **B10**: IR (KBr) 3379, 3147, 2958, 2932, 2871, 1705, 1630, 1465, 1276, 1196, 1100, 807, 780, 760 cm⁻¹; ¹H NMR (500 MHz, DMSO-d₆) δ 10.329 (s, 2 H, imide), 9.244 (s, 2 H, amide), 8.7529 (s, 1 H, amine), 8.306 (d, 2 H, J = 8.5 Hz), 8.302 (s, 2 H), 8.142 (apparent d, 1 H, J = 8 Hz), 8.043 (s, 1 H), 7.985 (m, 3 H), 7.807 (d, 2 H, J = 8.5 Hz), 7.65 (m, 1 H), 7.59 (m, 1 H), 7.511 (dd, 2 H, J = 8.8, 2.0 Hz), 7.386 (d, 2 H, J = 8.8 Hz), 4.46 (m, 2 H), 4.32 (m, 2 H), 3.89 (m, 2 H), 3.44

(m, 4 H), 3.00 (m, 1 H), 2.673 (d, 4 H, $J = 13.4$ Hz), 2.22 (m, 2 H), 2.0 (m, 4 H), 1.78 (m, 6 H), 1.51 (m, 4 H), 1.4-1 (m, 22 H), 1.013 (s, 9 H), 0.875 (t, 12 H, $J = 6.8$ Hz), 0.801 (t, 6 H, $J = 7.1$ Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for $C_{75}H_{91}N_8O_{10}$ (M-Cl), 1263.6858; found, 1263.6860.

N-(3,5-Bis(methoxycarbonyl)phenyl)carbazole (**B11**)

Thionyl chloride (55 mL) was added to a solution of 5-iodoisophthalic acid⁹³ (23.34 g, 79.9 mmol) in methanol (500 mL) on ice. After stirring for 22 h at room temperature, the mixture was refluxed for 9 h. Evaporation of solvent under reduced pressure afforded quantitative yield (25.58 g) of diester **B11**: mp 100.9-102.7 °C; IR (KBr) 3084, 2951, 1717, 1439, 1295, 1245, 996, 746 cm^{-1} ; 1H NMR (500 MHz, $CDCl_3$) δ 8.630 (t, 1 H, $J = 1.5$ Hz), 8.546 (d, 2 H, $J = 1.5$ Hz), 3.951 (s, 6 H).

N-(3,5-Bis(methoxycarbonyl)phenyl)carbazole (**B12**)

The iodoisophthaloyl diester **B11** (18.3 g, 57.17 mmol) was melted in an oil bath (110 °C). Carbazole (9.5545 g, 0.9995 equiv) and Cu_2O (25.55 g, 3.123 equivalents) were added with xylenes (2 mL). The reaction mixture was stirred at 170-175 °C for 48 h. The reaction mixture was cooled and extracted with benzene (3 x 50 mL) by refluxing, cooling and filtering. Flash chromatography (benzene) followed by recrystallization from refluxing MeOH (100 mL) and EtOAc (450 mL) gave 10.39 g (51%) of pure solid **B12**: mp 202.4-203.1 °C; IR (KBr) 2948, 1718, 1602, 1460, 1373, 1246, 750 cm^{-1} ; 1H NMR (300 MHz, $CDCl_3$) δ 8.774 (t, 1 H, $J = 1.5$ Hz), 8.458 (d, 2 H, $J = 1.5$ Hz), 8.154 (d, 2 H, $J = 7.8$ Hz), 7.437 (m, 2 H), 7.38 (m, 2 H), 7.326 (m, 2 H), 3.986 (s, 6 H); HRMS (EI) calcd for $C_{22}H_{17}N_1O_4$, 359.1158; found, 359.1155.

N-(3,5-Bis(methoxycarbonyl)phenyl)-3,6-dinitrocarbazole (**B13**)

Carbazole **B12** (2.15 g, 5.98 mmol) was added portion-wise over 35 min to $Cu(NO_3)_2 \cdot 2.5H_2O$ (4.57 g, 3.29 equiv) in HOAc (20 mL) and Ac_2O (40 mL) at 35 °C in an oil bath. After stirring for 17 h under a drying tube, the reaction mixture was added to 150 mL water, triturated for 30 min, filtered, and sucked dry to give dinitrocarbazole **B13** as an impure mixture which resisted purification. The estimated purity is 80%: 1H NMR (300 MHz, d_7 -DMF) δ 9.726 (d, 2 H, $J = 1.8$ Hz), 8.772 (t, 1 H, $J = 1.5$ Hz), 8.649 (d, 2 H, $J = 1.5$ Hz), 8.488 (dd, 2 H, $J = 9.1, 2.5$ Hz), 7.745 (d, 2 H, $J = 9.0$ Hz), 4.011 (s, 6 H).

N-(3,5-Bis(methoxycarbonyl)phenyl)-3,6-diaminocarbazole (**B14**)

Dinitro **B13** (955 mg, 2.125 mmol) was hydrogenated at balloon pressure in THF (166 mL) with 10% Pd-C catalyst (23 wt%) for 10 h. Filtration through Celite, followed by evaporation of solvent under reduced pressure gave an unstable yellow solid which was carried on immediately without further purification: 1H NMR (300 MHz, $CDCl_3$) δ 8.67 (t, 1 H), 8.41 (d, 2 H), 7.33 (d, 2 H), 7.22 (d, 2 H), 6.82 (dd, 2 H), 4.00 (s, 6 H).

N-(3,5-Bis(methoxycarbonyl)phenyl)-3,6-bis(((cis,cis-2,4-dioxo-1,5,7-tripropyl-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino)carbazole (**B15**)

Diamine **B14** (827 mg, 2.125 mmol) and tripropyl Kemp's imide acyl chloride²² (1.450 g, 1.995 equiv) with catalytic DMAP (11 mg) were refluxed in pyridine (50 mL) under argon for 10 h. The solvent was removed under reduced pressure and the residue taken up in CHCl₃ and EtOAc. The solution was washed with 1 N HCl and brine and dried over Na₂SO₄. Flash chromatography (10% EtOAc/CHCl₃ and 5% MeOH/CHCl₃) gave a yellow foam which was precipitated as a yellow solid (619 mg, 29%) by trituration in MeOH (4 mL): mp 203^{dec} °C; IR (KBr) 3379, 2958, 1728, 1701, 1467, 1247 cm⁻¹; ¹H NMR (500 MHz, DMSO-d₆) δ 10.373 (s, 2 H, imide), 9.258 (s, 2 H, amide), 8.535 (t, 1 H, J = 1.4 Hz), 8.333 (d, 2 H, J = 1.5 Hz), 8.302 (d, 2 H, J = 1.8 Hz), 7.513 (dd, 2 H, J = 8.8, 1.9 Hz), 7.357 (d, 2 H, J = 8.8 Hz), 3.932 (s, 6 H), 2.675 (d, 4 H, J = 14.0 Hz), 2.029 (d, 2 H, J = 12.5 Hz), 1.79 (m, 4 H), 1.51 (m, 4 H), 1.29 (m, 14 H), 1.15 (m, 8 H), 0.873 (t, 12 H, J = 7.0 Hz), 0.807 (t, 6 H, J = 7.2 Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₅₈H₇₄N₅O₁₀ (M+H), 1000.5436; found, 1000.5426.

5.2 Titrations

Association constants were calculated by non-linear least-squares regression fit of the data to the 1:1 binding isotherm using Systat 5.2.³⁵ While the final chemical shift and association constant were treated as variables; the experimental value for the initial chemical shift was used.

¹H monitored

¹H titrations were performed at 298 K either at 500 MHz on a Varian VXR-500 instrument with a 5 mm ¹H/¹⁹F indirect detection probe and an Oxford VTC-4 temperature controller. A 1 mM solution of receptor **B9** in CD₃OD (99.8%-D, CIL, 700 μL) was titrated with a 10 mM solution of cyclic AMP in CD₃OD while the chemical shift of carbazole C⁴H was monitored. The total chemical shift changes were 0.19 ppm (3',5'-cAMP) and 0.22 ppm (2',3'-cAMP) downfield; 90-94% of the saturation curves were observed.

³¹P monitored

³¹P titrations were performed at 121.4 MHz on a Varian UN-300 instrument equipped with a 5 mm ¹H/¹⁹F broad band probe and an Oxford VTC-4 temperature controller. A solution of 2',3'-cAMP (1 mM), 3',5'-cAMP (1 mM), and NaBr (23 mM, I=25 mM) in CD₃OD (99.8%-D, CIL, 700 μL) was treated with a solution of **B9** (10.4 mM), 2',3'-cAMP (1 mM), 3',5'-cAMP (1 mM), and NaBr (13 mM, I=25.7 mM) in CD₃OD while the chemical shift of the phosphate group was monitored. The total chemical shift change was 0.321 ppm upfield for 2',3'-cAMP and 1.076 ppm upfield for 3',5'-cAMP; 61-73% of the binding curves were observed. Triphenylphosphine was used as an internal shift standard.

5.3 Modeling

Molecular modeling was performed on Silicon Graphics Personal Iris (4D25G+ or 4D30G+) machines running MacroModel 3.5X.⁷⁰ The AMBER* force field was used with GB/SA⁷³ continuum water solvation. Stochastic dynamics⁷¹ simulations were run for 65 ps at 300 K with a 1.5-fs timestep using the SHAKE algorithm to constrain all bond lengths. The distance between the imide nitrogen's and the correspondingly-bonded adenine nitrogen was restrained to 2.8 ± 0.5 Å with a 100 kJ/mol potential. After a 5-ps induction period, 300 structures were sampled for the duration of the simulation.

Following the simulation, the sampled structures were minimized with the truncated Newton conjugate gradient algorithm⁹⁴ (TNCG) to the default gradient (<0.05 kJ/mol-Å), using the MULT routine in MacroModel with 0.2 Å as the maximum allowed deviation for comparison of all heavy atoms. For the 2',3'-cAMP-bound receptor, 57 unique conformations were found within 31 kJ/mol of the global minimum. There were 33 conformations within 10 kJ/mol of the global minimum. The 3',5'-cAMP-bound receptor showed a total of 4 unique conformations, all within 4 kJ/mol of the global minimum.

5.4 Transport

Transport studies were performed by Cecilia Andreu using a U-tube with dichloroethane separating the aqueous source and receiving solutions (2 mL). The source phase was a solution of the nucleotide and the receiving phase was an NaCl solution, 10 mM, unless otherwise noted. The surface area of contact is 1.568 cm² between aqueous and organic phases.

The rate of transport was followed by monitoring the concentration of nucleotide in aliquots from both aqueous phases by UV/visible spectroscopy. Experiments were performed at least in duplicate.

C. SELF-REPLICATING MOLECULES: A SECOND GENERATION

1. BACKGROUND

Reproduction is essential to life.⁹⁵ In prebiotic times, reproduction was merely the replication of stable molecular assemblies drawn from the amino acids,⁹⁶ carbohydrates,⁹⁷⁻⁹⁹ heterocyclic bases,¹⁰⁰ and enzyme cofactors¹⁰¹⁻¹⁰³ expected under the conditions of early Earth. As such, the self-replication of molecular structures lies at the heart of the molecular origins of life.

Molecular replication is simply an autocatalytic reaction where the product of a chemical transformation acts to direct or catalyze that transformation. This typically manifests itself as a template effect. That is, the reaction product brings the reactants together to accelerate the reaction through non-covalent interactions. This has been implemented in the replication of physical structures and self-complementary molecules.

The replication of physical structures is exemplified by the autocatalytic generation of micelles in a system developed by Luisi.¹⁰⁴ Ethyl caprylate is hydrolyzed at the aqueous-organic interface in a two-phase system to generate sodium caprylate which forms micelles in the aqueous layer. The hydrophobic ester is solubilized by the micelles; this increases its exposure to base and accelerates hydrolysis by almost three orders of magnitude. After an induction period, exponential growth in the concentration of sodium caprylate occurs, a characteristic of autocatalytic reactions. This is an interesting system because it suggests a mechanism for the development of cells outside of metabolism or information transfer.

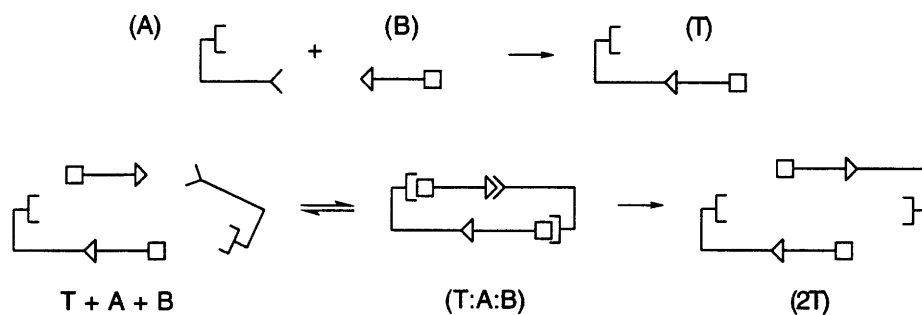


Figure C.1.1. Schematic of self-complementary template-based autocatalysis.

Self-complementary autocatalysis relies on an efficient reaction within a termolecular complex. As shown in Figure C.1.1 (from Edward Winter), two complementary components of a replicating structure (A and B) react in an intermolecular fashion to form template (T). Due to the now *self*-complementary nature of the template, two additional units of A and B form a termolecular complex with the template. Intracomplex reaction forms a dimer that dissociates to repeat the process. Ideally, some form of catalysis occurs in the termolecular complex, but, typically, it is simply the reduction in entropy of bringing two

reagents close together in a complex that enhances the reaction rate.

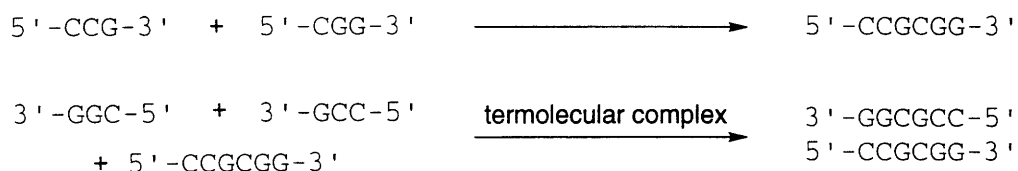


Figure C.1.2. Self-replication of DNA trimers.

The first self-replicating model system developed used self-complementary DNA as the basis for the autocatalytic reaction. Trideoxynucleotides CCG and CGG were coupled in the presence of a water-soluble condensation agent to form the self-complementary hexadeoxynucleotide by an autocatalytic process (Figure C.1.2).¹⁰⁵ Improvements involved the use of 5'-deoxy-5'-aminodeoxynucleotides to accelerate nucleophilic attack to increase autocatalytic efficiency so that parabolic growth in the hexadeoxynucleotide concentration could be seen, reflecting the exponential nature of an autocatalytic process.¹⁰⁶ Recently, a self-replicating system utilizing three DNA components has been developed to investigate the role of competition in the self-replication of a DNA model system.¹⁰⁷

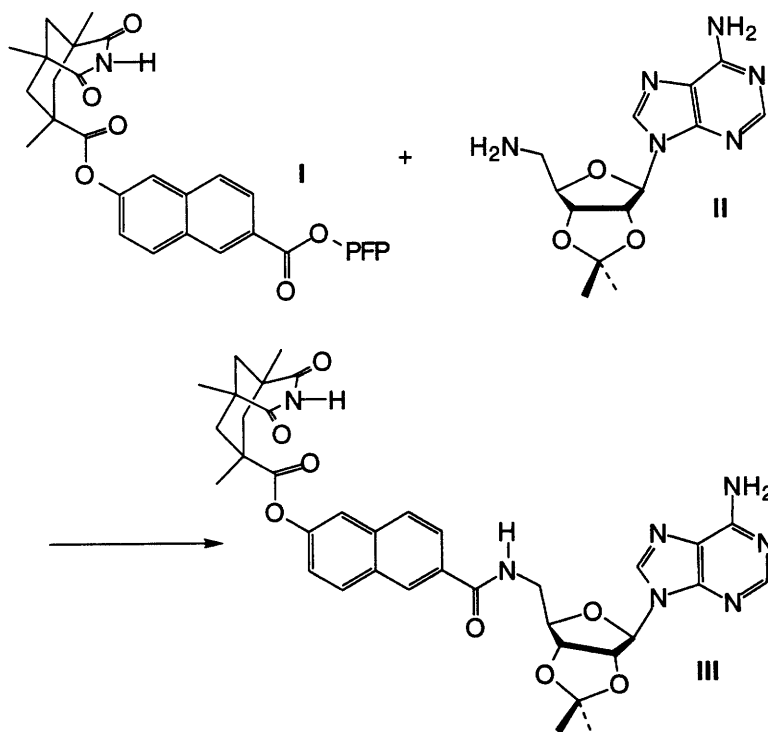


Figure C.1.3. The first abiotic self-replicating system.

The first self-replicating system designed from extrabiological structures was reported by this laboratory.¹⁰⁸ The hydrogen-bonding recognition of adenine and a Kemp's triacid imide provided the driving force for supramolecular recognition. The autocatalytic reaction is diagrammed in Figure C.1.3 as the reaction

between imide-bearing naphthoyl ester **I** and aminoadenosine **II** to generate the self-complementary template **III**. While the catalytic efficiency within the termolecular complex is insufficient to observe exponential growth of the template, the autocatalytic nature of the reaction is evident from the rate acceleration caused by seeding the reaction with its product.

A number of control studies were performed to ascertain that the template effect is responsible for the autocatalytic rate acceleration (Figure C.1.4).¹⁰⁹ The reactions were monitored by following the initial rate of product formation through HPLC analysis of a 2 mM chloroform solution of starting materials with 0.5 equiv (1 mM) of the 'catalyst' added in the presence of excess triethylamine (72 mM). Under these conditions, addition of template produced a 50% increase in the reaction. The N-methylated imide **IV** presents all of the functionality of the template, but precludes binding through imide-adenine molecular recognition. As a result, no rate enhancement, within 2% experimental error, results when this is added to the reaction. Furthermore, an abbreviated analog, naphthoylated adenosine **V**, is also ineffective and even exhibits marginal inhibition of the reaction. It has been reported that primary, secondary, and tertiary amides are capable of catalyzing aminoacylation reactions by acting as O-acylation catalysts.¹¹⁰ The amides presented in the context of these systems do not appear to act in this fashion.

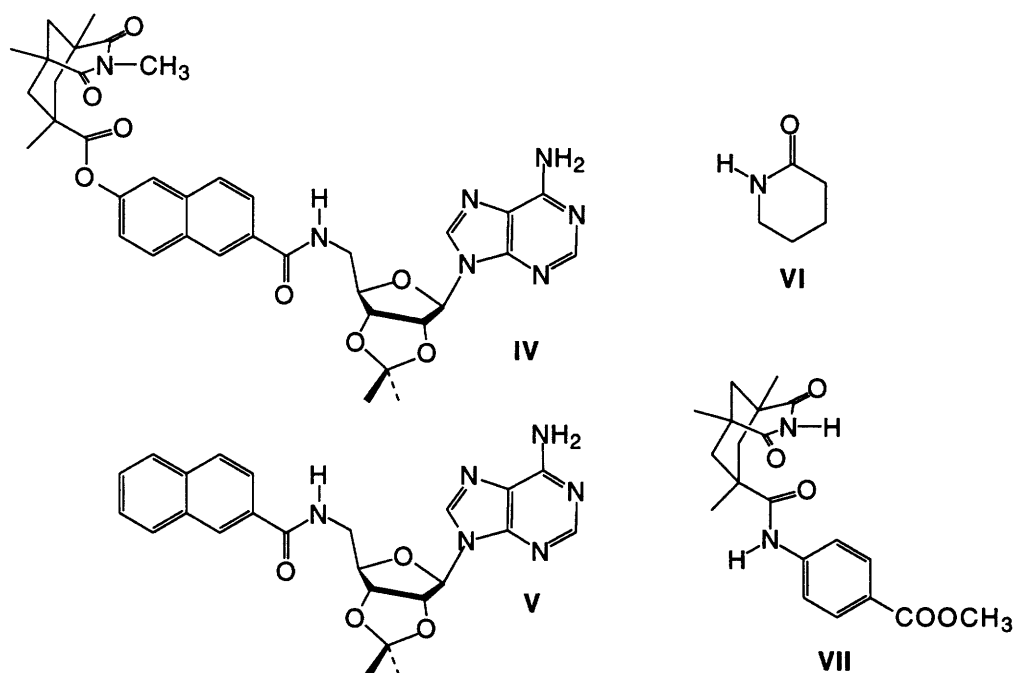


Figure C.1.4. Controls for catalysis by self-replication.

Under these same conditions, valerolactam, **VI**, produced a 16% rate enhancement. This is not surprising. Menger has established that in aprotic organic solvents the breakdown of the tetrahedral intermediate is the rate-determining step in the acylation of amines.¹¹¹ Valerolactam can interact with the tetrahedral intermediate to facilitate the required proton transfers for product release. The role of

bifunctional catalysts in acylation reactions¹¹² and glucose mutarotation¹¹³ is well known. Figure C.1.5 depicts the role we envisage for the lactam in catalysis. It is likely that primary or *cis*-amides could catalyze the aminoacylation reaction in much the same way.¹¹⁴ Catalysis by valerolactam through an O-acylation mechanism is also a distinct possibility.¹¹⁰ Clearly, though, the amide in autocatalytic template **III** does not at all contribute to catalysis outside of complexation. Interestingly, imide **VII** is also ineffective as a catalyst, despite presenting functionality in much the same way as valerolactam. The aromatic group, roughly parallel to the imide, apparently lies too close to allow interaction between the imide and tetrahedral intermediate. If the operative mechanism is catalysis through O-acylation, then the lower basicity of the carbonyl in the imide group may also be the reason for this.

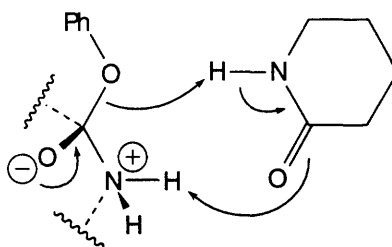


Figure C.1.5. Schematic of lactam catalysis of tetrahedral intermediate breakdown.

It may be possible that in the termolecular complex between **I**, **II**, and **III**, the amide on **III** exerts some catalytic effect on the tetrahedral intermediate. It has been postulated that the appropriate placement of electrostatic fields in enzyme active sites is, in large part, responsible for enzymatic catalysis.¹¹⁵ This type of effect occurs in this system, but the control studies performed make it clear that this could play a role *only within the termolecular complex*.

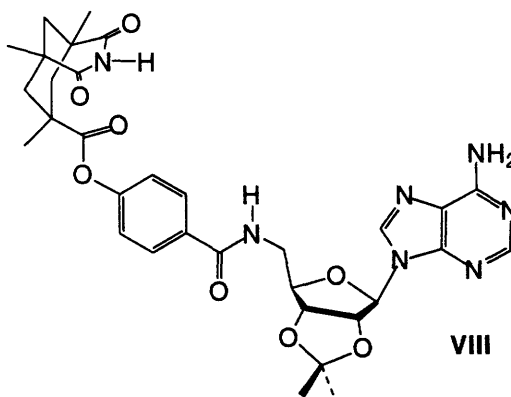


Figure C.1.6. Control for fit-constraints in self-replication.

The phenyl-based imide, **VIII**, was used as a probe of the strictness of fit in the termolecular complex. The molecule dimerizes in solution with an association constant of approximately 800 M^{-1} , much similar to the naphthyl-based template.¹¹⁶ Molecular modeling suggested that, while phenyl **VII** and naphthyl **III** do not associate well as a mixed 'dimer', the phenyl template may form a complex with the tetrahedral

intermediate that forms between amine **II** and ester **I**. This could, in fact, become the basis for efficient catalysis, since complications from product inhibition may be reduced. However, within the limits of experimental error, **VII** showed *no catalysis* at all under our conditions.

One of the complications of the original naphthoyl-based self-replicating system is that coupling between amine and ester can occur within a complex of the two precursors, to form a *cis*-amide, which isomerizes to the *trans*-amide to form active template. This pathway is accessible whether the adenosine is bound in the Watson-Crick or Hoogsteen mode. Figure C.1.7 shows end-on views of the Hoogsteen (left) and Watson-Crick (right) complexes with the amine in close approach to the ester. This pathway dominates over autocatalysis for template formation.¹¹⁶ The first step towards restricting this involved changing the spacer element to a biphenyl,¹¹⁷ where the ester and amine are moved away from each other. Intracomplex reaction can only occur if the adenosine is bound in the Watson-Crick mode (Figure C.1.8). By reducing the amount of preassociative catalysis, the effect of the autocatalytic pathway can be seen as a slight sigmoidality in the initial rate of growth of product. This has been exploited for the study of competition and mutation using N⁶-Cbz derivatives of adenosine.¹¹⁸

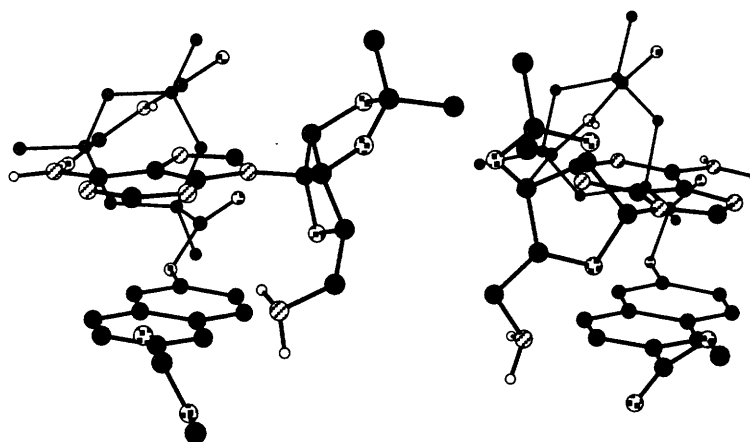


Figure C.1.7. Intracomplex reaction in the naphthoyl self-replicating system.

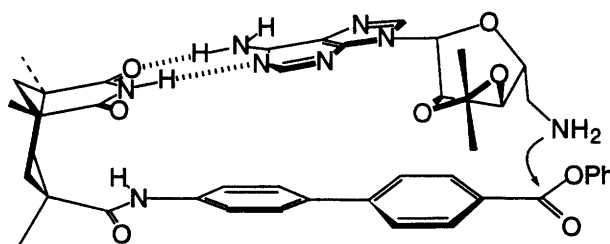


Figure C.1.8. Intracomplex reaction in the biphenyl self-replicating system.

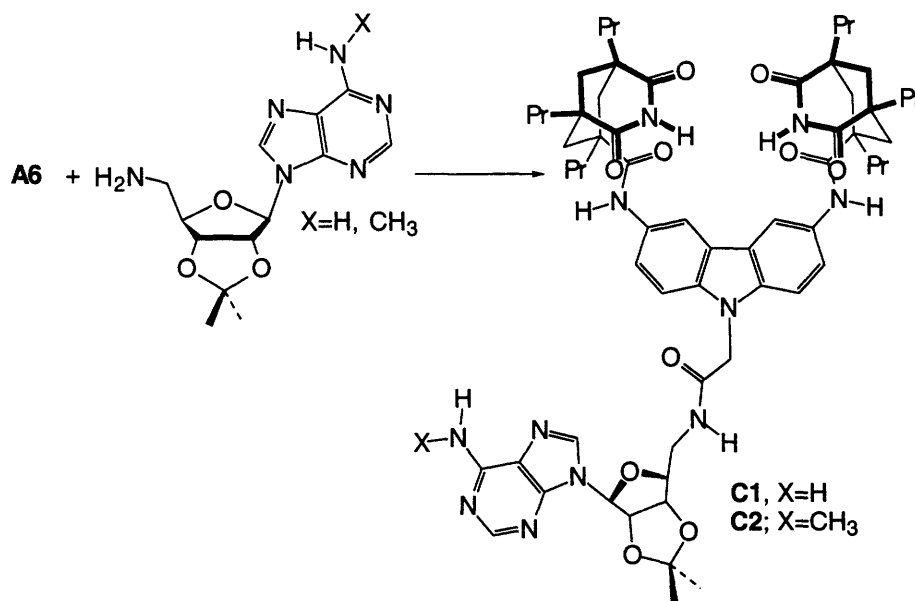
In order to optimize these self-replicating systems, then, the preassociative bimolecular pathway must be restrained. The spacer element in the receptor must be sufficiently long to separate the amine and ester in a bimolecular complex, regardless of the binding mode of adenine. Extending abiotic self-replication into more polar solvents, and eventually water, would allow these systems to be evaluated on the same

playing field as biologically-derived replicating systems. In order for this to be realized, the binding affinity of the receptors to adenosine must be significantly enhanced so that molecular recognition occurs in more competitive solvents.

Both of these criteria can be fulfilled by the diaminocarbazole-based diimide module. Not only is the binding affinity for adenosine derivatives extremely high, but triplex-like chelation unambiguously fixes the conformation of the bound adenine.

2. SELF-COMPLEMENTARY MOLECULES

A number of backbone geometries were explored in the process of developing a self-replicating system based on the 3,6-diaminocarbazole spacer element. The first of these is the self-complementary derivative of the original carbazole-diimide (**A5**). Condensation of 5'-amino-5'-deoxy-2',3'-isopropylideneadenosine with acid **A6** (Scheme C.2.1) provides 'template' **C1**.



Scheme C.2.1. Synthesis of methylscorpion template.

The proton NMR spectrum of this molecule in CDCl₃ indicates that the diimide binding site is completely occupied by adenine. As has been seen previously for these systems (Section A.2), the slow exchange in these systems cause broadening of the resonances. At 25 °C, the ¹H NMR spectrum of a 10 mM solution of **C1** at 500 MHz shows two broad imide signals centered on 13.25 ppm (Figure E.7d). Variable temperature spectra from 0 °C to 40 °C (Figure E.7) show that the two imide peaks coalesce at 40 °C. The rate of exchange between the two imides at this temperature can be calculated from the maximum chemical shift difference between the two resonances. This is estimated as 93 Hz, from the spectrum at 0 °C where the peaks are completely separated. Using standard dynamic NMR spectroscopy theory as

outlined in Section A.2, at the coalescence temperature, the rate of interconversion, k_r , is approximately 207 s^{-1} and the barrier to interconversion, ΔG^\ddagger , is 15 kcal/mol. This barrier is somewhat higher (1.4 kcal/mol) than for the complex of **A24** and 9-ethyladenine.

No shift in the imide signal can be seen on dilution of this solution to 0.1 mM, but there appears to be a broadening of the imides at high dilution (Figure E.8). This may be an artifact of the low signal-to-noise ratio.

There are two possible interpretations for the dilution and variable temperature NMR studies. Though it is clear that the molecule binds tightly to itself, it is not obvious whether this is in an inter- or intramolecular sense. On the basis of molecular modeling, the shape of the backbone should allow the adenine to fold back and bind within the diimide pocket with some distortion (Figure C.2.1). The adenine and carbazole aromatic groups deviate significantly from coplanarity and the overlap between the two groups is poor, forcing the bicyclic imide groups 'back' in order for binding to occur.

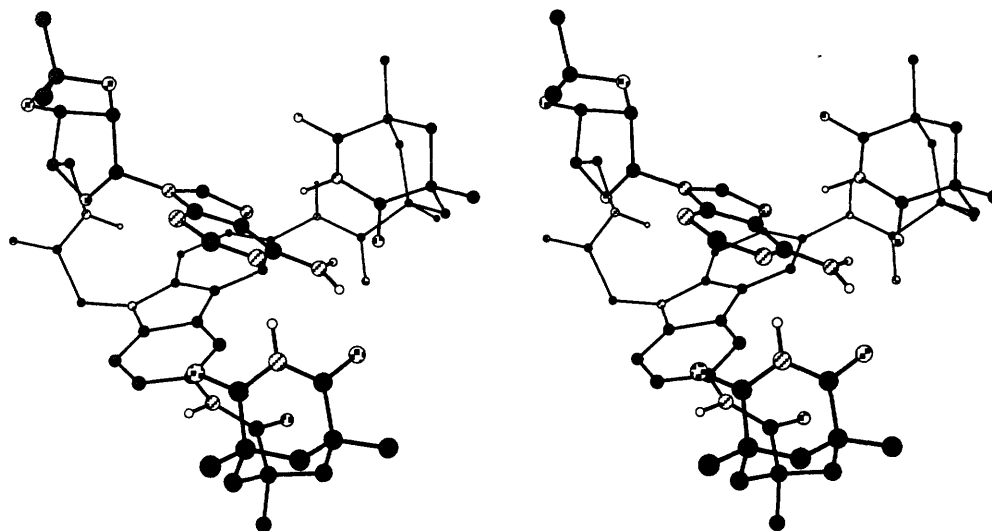


Figure C.2.1. Predicted geometry of the intramolecular conformation of **C1** (stereoview).

On the other hand, dimerization is extremely strong in this solvent. An estimate of this can be made, given the binding constant of a single diimide-adenosine association event. If there is no significant strain in the dimerized structure (a key assumption), then the dimerization interaction energy will be at least the sum of two individual association energies. In fact, the second binding event in dimerization should be significantly more favorable than the first since the translational and rotational entropy between the two molecules has already been overcome. However, if this increase in interaction energy, which has been estimated to be worth 2-11 kcal/mol,¹¹⁹ is neglected, then,

$$K_{\text{dim}} \approx e^{-2\Delta G_a/RT} = \left(e^{-\Delta G_a/RT}\right)^2 = K_a^2 \quad (\text{C.2.1})$$

The individual components, 5'-amino-5'-deoxy-2',3'-isopropylideneadenosine and methyl ester **A5** would be expected to have an association constant on the order of 50000 M^{-1} (see Section A.3). This would lead

to an estimated dimerization constant of $2.5 \times 10^9 \text{ M}^{-1}$. The following sequence (Equations C.2.2 to C.2.8) illustrates the method by which the chemical shift of the imide can be simulated. The dimerization constant, K_d , is defined in C.2.3 for the equilibrium represented in C.2.2.



$$K_d = \frac{[A_2]}{[A]^2} \quad (\text{C.2.3})$$

The total concentration of the self-complementary molecule is A_t and:

$$A_t = [A] + 2[A_2] \quad (\text{C.2.4})$$

By combining Equations C.2.3 and C.2.4 and solving for $[A_2]$, we get:

$$[A_2] = \frac{1 + 4A_t K_d - \sqrt{1 + 8A_t K_d}}{8K_d} \quad (\text{C.2.5})$$

Since the amount of A_t as monomer is $[A]/A_t$, then the amount as dimer is $2[A_2]/A_t$, from C.2.4. Therefore,

$$\frac{2[A_2]}{A_t} = \frac{1 + 4A_t K_d - \sqrt{1 + 8A_t K_d}}{4A_t K_d} \quad (\text{C.2.6})$$

And the amount in solution as monomer is,

$$\frac{[A]}{A_t} = 1 - \frac{1 + 4A_t K_d - \sqrt{1 + 8A_t K_d}}{4A_t K_d} \quad (\text{C.2.7})$$

Using these two expressions, C.2.6 and C.2.7, the chemical shift, δ , of the imide can be estimated from the chemical shifts of the free, δ_f , and fully bound, δ_b , species.

$$\delta = \frac{[A]}{A_t} \delta_f + \frac{2[A_2]}{A_t} \delta_b \quad (\text{C.2.8})$$

If we use values of 13.2 ppm for the shift of bound imide and 10.2 ppm for shift of 'free' imide, which intramolecularly bound to the other imide when not complexed (see section A.2), and a dimerization constant of $2.5 \times 10^9 \text{ M}^{-1}$, then the variation of chemical shift with concentration can be calculated and is shown in Figure C.2.2.

This simulated data clearly shows that a binding constant of this magnitude will require concentrations as low as 10 pM to see a 0.01 ppm change in the imide chemical shift. This is far below the sensitivity of NMR spectroscopy, therefore dilution studies should be considered of only limited use in discerning dimerization from intramolecular association when binding constants are large.

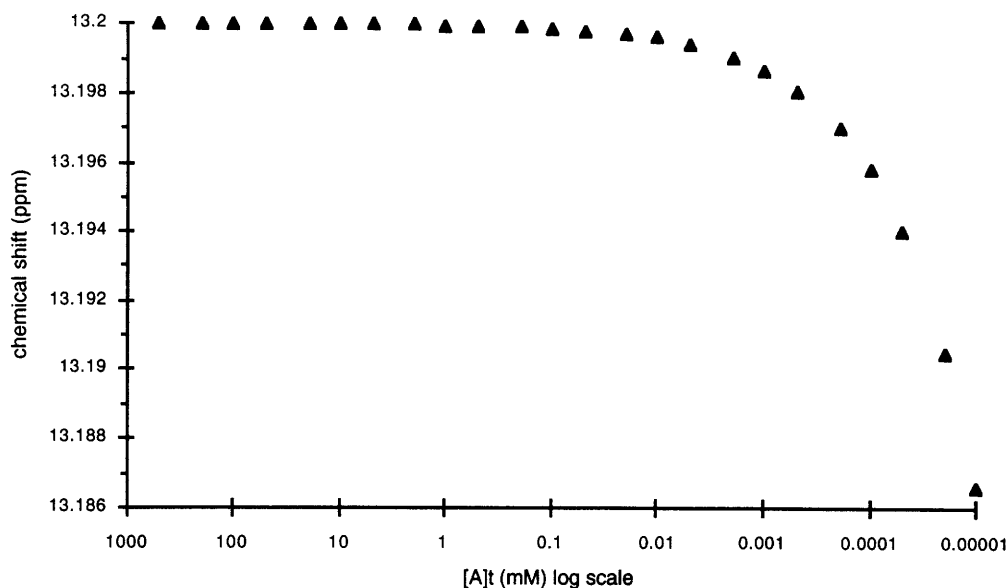


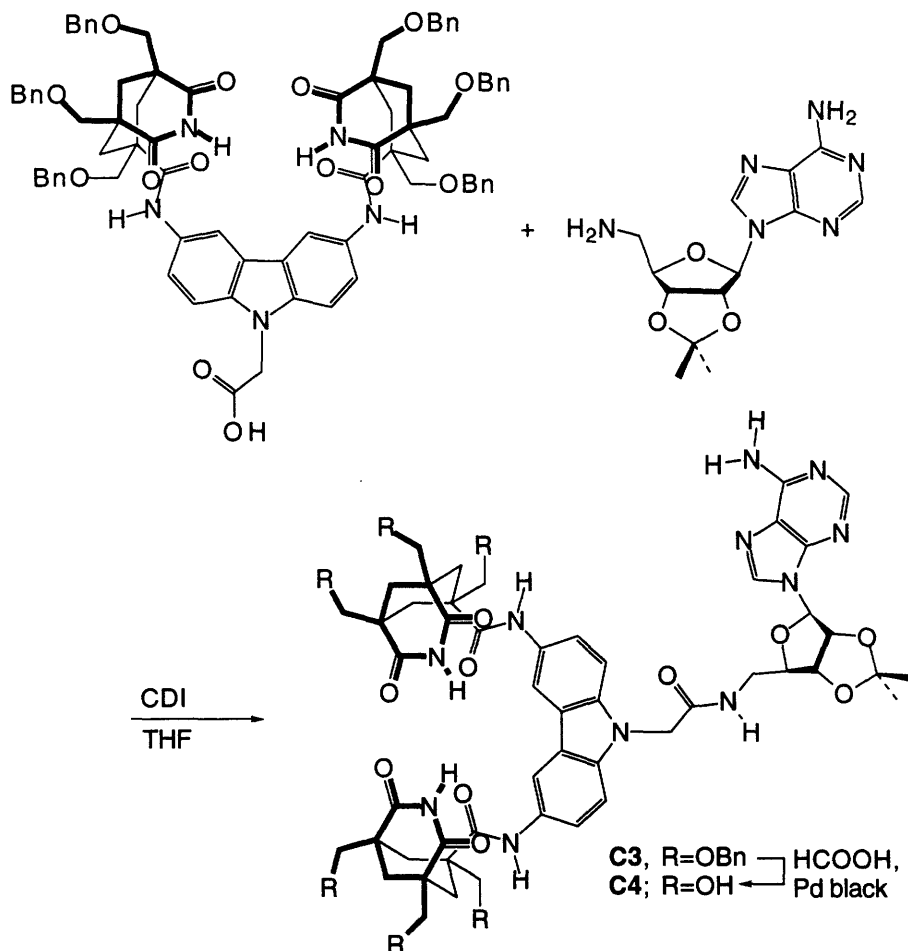
Figure C.2.2. Simulation of imide chemical shift with concentration.

Control **C2** was synthesized to ascertain the association stoichiometry of **C1**. The methyl group on the exocyclic amine of adenosine prevents simultaneous complexation in the Watson-Crick and Hoogsteen binding modes and reduces binding to a single imide-adenine contact, an association constant of 165 M^{-1} (Section A.3). The dimerization of such a species would then be estimated to be approximately 27000 M^{-1} . By applying the analysis of equations C.2.2 to C.2.8, the imide proton would be expected to move 0.15 ppm upon dilution to 0.1 mM if the system was dimerized. The ^1H NMR spectrum showed no change in the solitary imide resonance at 11.7 ppm on dilution from 11.25-0.085 mM in CDCl_3 (Figure E.9). This chemical shift is the average of the bound (13.2 ppm) and uncomplexed (10.2 ppm) imide shifts, indicating rapid imide exchange between bound and unbound states. The insensitivity to dilution suggests that the self-complementary molecule based on the framework of **A5** associates through an intramolecular process.

Additional support for intramolecular association comes from vapor pressure osmometry (VPO) measurements performed by Ivan Huc. The apparent molecular weight of a 3 mM solution of **C1** in chloroform is 833 amu, 72% of the actual value (1154.6 amu), using a sample of **B5** as a standard. The accuracy of the VPO method is notoriously poor as the technique works optimally for non-associating molecules. Systems with polar functionality tend to give higher-than-theoretical molecular weights due to nonspecific aggregation.¹²⁰ The slight mismatch of the standard and sample contributes to the underestimation of the weight. (it has been seen that on the basis of a single phenyl ring, **B5** and **C13** generate different standard lines¹²¹).

A water-soluble version of **C1** was produced by coupling the benzyloxymethyl derivative of **A6**, made by

Kazuya Kobi, with 5'-amino-5'-deoxy-2',3'-isopropylideneadenosine (Scheme C.2.2). The organic-soluble precursor, **C3**, behaved similarly to the propyl derivative **C1** by NMR spectroscopy with a single imide peak at 13.43 ppm from 0.13-9.94 mM in CDCl₃. It is significant that only one imide signal is observed at 25 °C. The two hydrogen-bonding modes are in more rapid exchange with this molecule than in the propyl derivative. This increased exchange rate may arise from accelerated complex-opening due to the extra bulk of the benzyloxymethyl groups. Hydrogenolysis with formic acid and palladium black catalyst¹²² produces the water-soluble self-complementary **C4**.

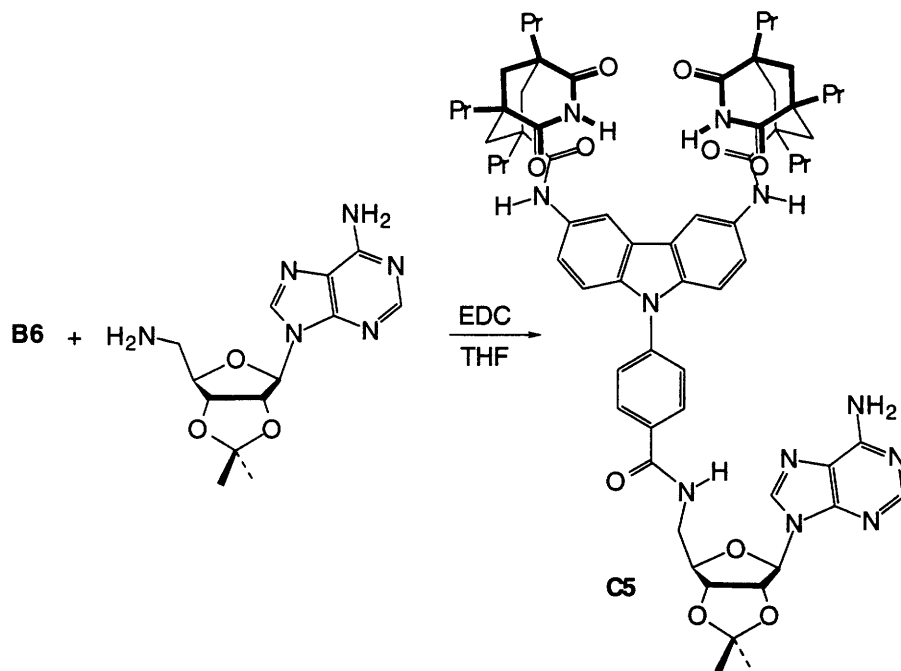


Scheme C.2.2. Synthesis of tris(hydroxymethyl) methylscorpion template.

Phenylscorpion **B5** was employed for the synthesis of a self-complementary scorpion, **C5**, that features the adenosine group more remotely positioned from the diimide cleft.

Due to the limited utility of dilution studies, none were performed, but a variable temperature ¹H NMR study indicated that **C5** exists in a very stable conformation (Figure E.10). In contrast to the molecule with a biphenyl spacer (see Section C.3.1), the two imide proton signals remain conspicuously sharp and separated even at 50 °C in CDCl₃. On cooling, these two resonances broaden slightly and, from -20 °C,

four small additional imide signals arise. With slightly broadening at low temperature and sharpening at high temperature, along with no coalescence in the range of accessible temperatures, the behavior of the split imide doublet is completely unlike either **C1** or **C15** (Section C.3).



T

Scheme C.2.3. Synthesis of phenylscorpion template.

The origin of this complex behavior is not clearly understood. There appear to be multiple species in solution, with one preferred and others forming at lower temperatures. In addition to the dimerization expected for a self-complementary molecule, molecular modeling indicates that it is possible for the molecule to fold shut if the amide bond is in a *cis* conformation. The predicted structure of the folded structure is shown in Figure C.2.3. It can be seen that the aromatic moieties retain coplanarity and four hydrogen bonds are formed. The Hoogsteen face of the adenine is bound in a bifurcated sense by rotation of one of the imides in relation to the plane of the carbazole.

The formation of a *cis*-amide in this system is not unlikely. The *trans*-amide conformation of *N*-methylacetamide is predicted to be only 2 kcal/mol more stable in solution than the *cis*-amide.¹²³ Despite this apparent low barrier, non-proline *cis*-amide bonds occur only infrequently (0.05%) in protein structures, though bias during structure solving may influence this somewhat.¹²⁴ The energetic cost of a secondary *cis*-amide was predicted to be 5 kcal/mol based on a recent protein mutagenesis study incorporating a glycylyl *cis*-amide in place of a proline residue.¹²⁵ Even so, the steric environment of the aromatic to adenosine amide bond in **C5** is much more forgiving than in peptides and the interaction energy associated with adenine-diimide binding, 6.4 kcal/mol for the phenylscorpion (see Section B.2), should be well in excess of the *cis-trans* energy difference. Moreover, the interaction energy would be

expected to be even stronger for intramolecular association, since intermolecular translational and rotational entropy need not be overcome. It is quite possible that the unfavorable conformational energy of a *cis*-amide bond is more than offset by the binding energy of diimide-adenine recognition. Unfortunately, infrared spectroscopy is unable to reveal any trace of the *cis*-amide, as the carbonyl region of the spectrum is dominated by other functionalities.

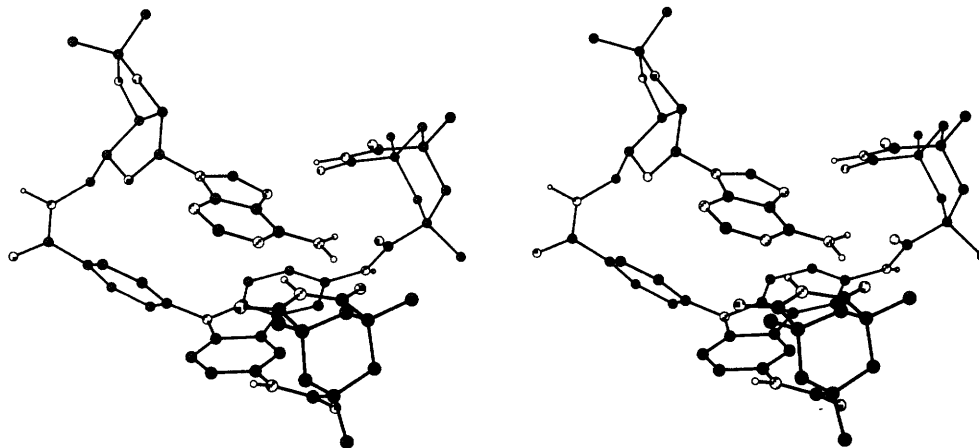


Figure C.2.3. Predicted structure of folded **C5** (stereoview).

The large changes in translation and rotational entropy inherent to intermolecular associations will, in general, cause them to be more sensitive to temperature than the corresponding intramolecular process. A likely interpretation of the NMR data is that an equilibrium exists in solution between the *cis*-amide intramolecularly-bound conformation and the dimeric species. At higher temperature, the intramolecular complex is preferred and maintains a tight fold, but as the temperature decreases, the equilibrium shifts to multimolecular aggregation.

Vapor pressure osmometry measurements only tentatively support this. The molecular weight determined, using **B5** as a standard, was 2000 amu. This is 164% of the expected value for a monomeric species (1216.5 amu). This is the same margin of error seen for the biphenylscorpion **C15** (Section C.3), a case with known binding stoichiometry. While this is not hard evidence, it can be used to suggest that the system is monomeric in solution. A contributing factor to the VPO error may be the nonspecific aggregation of the monomeric species through the *cis*-amide. The dimerization constant of a *cis*-amide (as exemplified by a β -lactam) is approximately $2\text{--}4\text{ M}^{-1}$.¹²⁶

Since the self-complementary scorpions have shown the ability to form folded structures, a molecule was designed with a backbone that would exist predominantly in a folded conformation with little distortion of the binding cleft of backbone.

The purpose of this was to simplify the formation of crystals of a complex between the diimide cleft and adenine. As the carbazole-diimide module is stubbornly non-crystalline, it was hoped that suitable crystals could be obtained for atomic structure determination by forming individually closed structures. Molecular

modeling of the target system indicates that the two aromatic moieties are coplanar with a *trans*-amide when the adenine and diimide cleft are intramolecularly-bound. As they are designed to snap shut, these molecules are, in a sense, molecular jackknives.

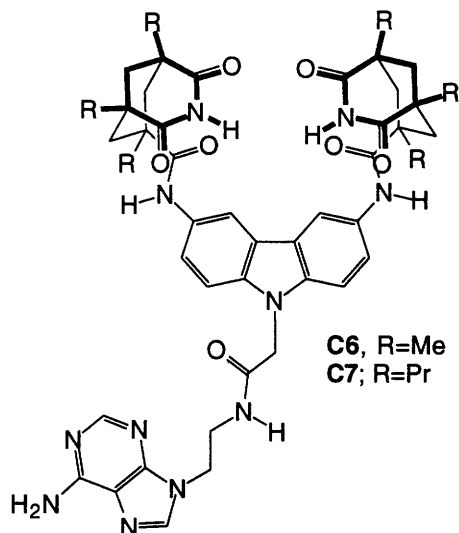


Figure C.2.4. A molecular jackknife.

As well, the methyl Kemp's triacid derivatives were employed to exploit their greater crystallinity. Unfortunately, all attempts to form crystalline material were unsuccessful, though the solubility of both **C6** and **C7** was found to be poor.

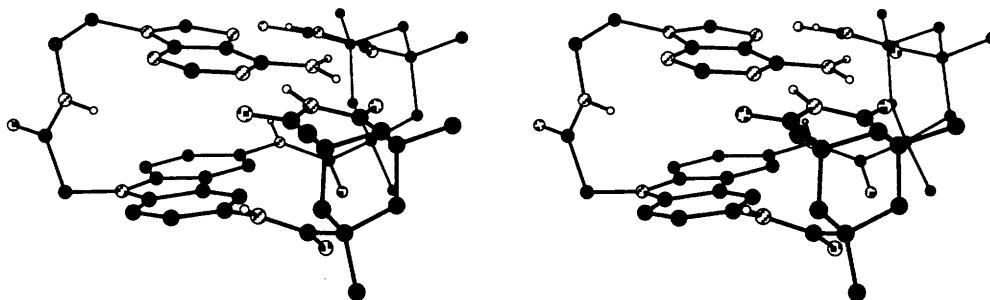


Figure C.2.5. Predicted structure of the molecular jackknife.

3. AUTOCATALYSIS

The intramolecular folding of self-complementary molecules reduces their application to autocatalytic processes in two ways: *i*) the folding reduces or removes the amount of molecule that can act as a template in an autocatalytic process and *ii*) if the molecule, as a whole, can fold, then the possibility of the preassociative bimolecular pathway occurring is significantly enhanced. The efficiency of the termolecular catalytic can easily be hidden if the intracomplex reaction dominates. In order to design an optimum autocatalytic system, then, the reactive centers within the complex should be held apart, unable to react with each other. That is, reaction should *only* occur in an intermolecular sense or in the termolecular

complex.

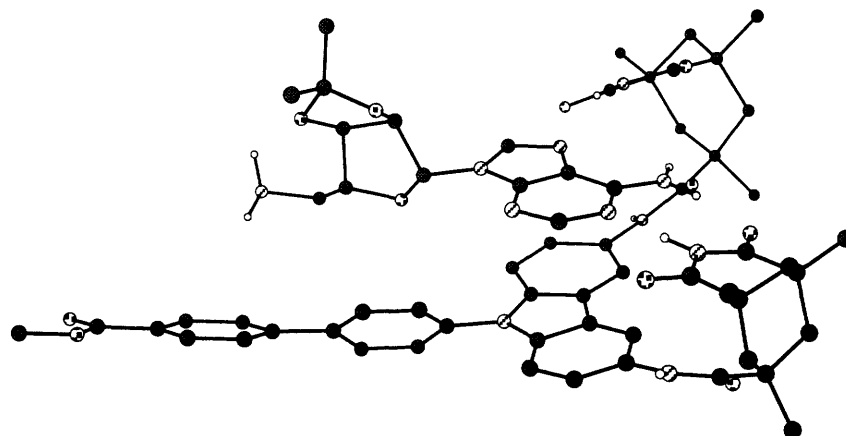


Figure C.3.1. The complex between the biphenylscorpion and aminoadenosine.

REACTION I:

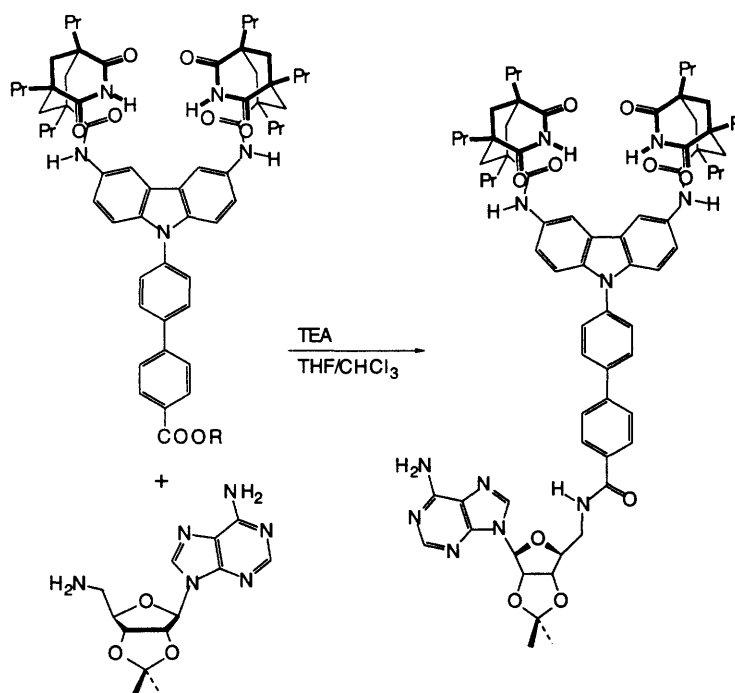
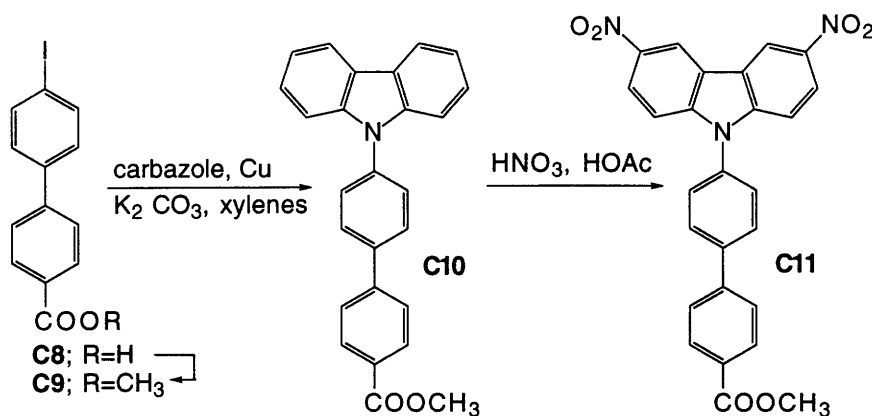


Figure C.3.2. The autocatalytic reaction.

The previous section showed that existing backbone structures are unable to separate the amine and reactive ester moieties within the complex. A biphenyl substituent on N^9 of the carbazole-diimide is ideal to enforce this separation. Figure C.3.1 shows the predicted geometry of a bimolecular complex between 5'-amino-5'-deoxy-2',3'-isopropylideneadenosine and biphenylscorpion **C13** with the amine and ester separated by a significant distance, 5.86 Å. This is a well-defined distance as there is only limited motion of diimide-bound adenine within the complex. Since the two reactive centers cannot approach each other, then reaction to form the template molecule must occur in an intermolecular sense.

This new architecture features both the high affinity of the carbazole-diimide module and elimination of the preassociative bimolecular pathway for the proposed autocatalytic reaction shown in Figure C.3.2. The self-complementary product has a backbone geometry that does not allow any intramolecularly folded species, regardless of the conformation of the amide bond.



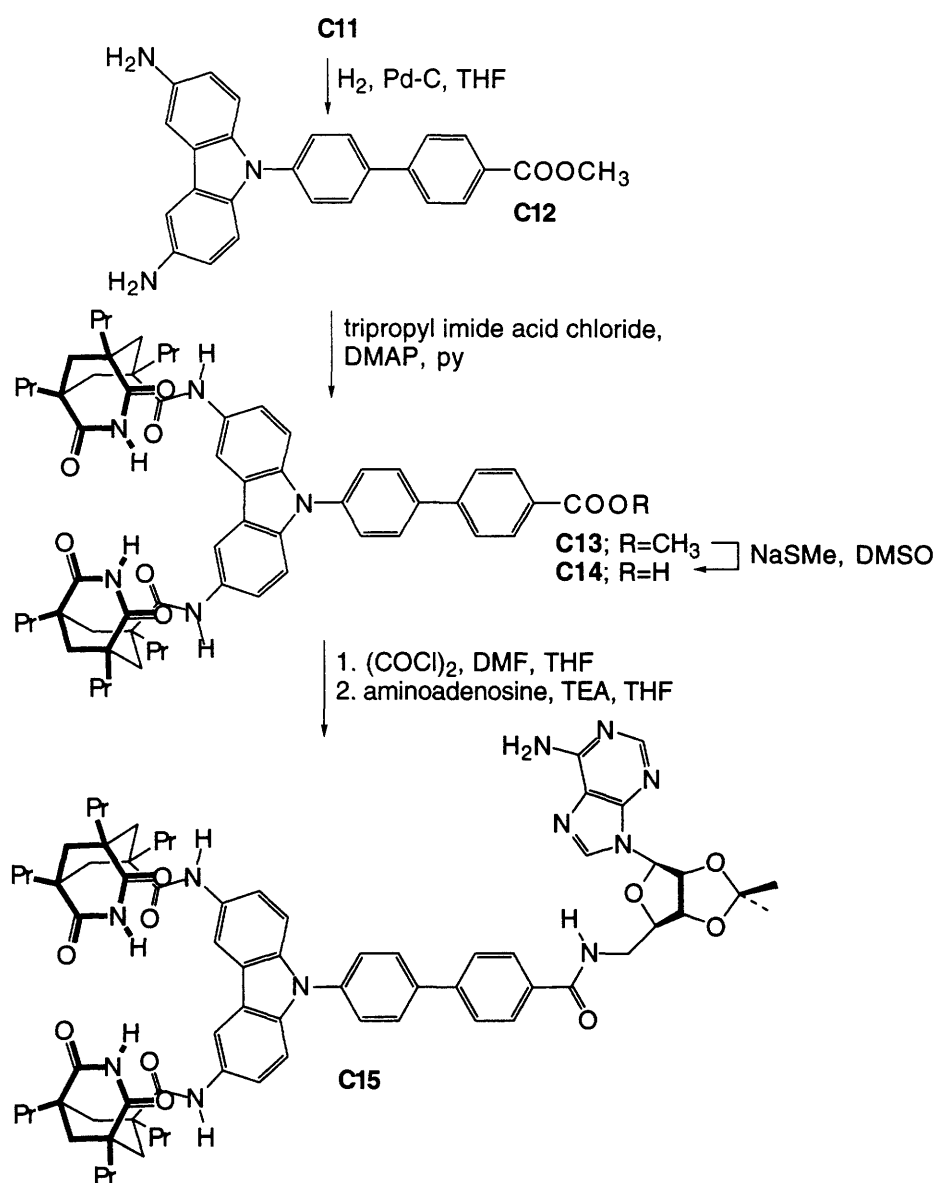
Scheme C.3.1. Synthesis of N-(4-(4'-(methoxycarbonyl)phenyl)phenyl)-3,6-dinitrocarbazole.

Synthesis of the biphenylscorpion proceeded through Ullman coupling⁸⁶ of iodobiphenyl derivative **C9** and carbazole followed by nitration to give the aromatic backbone **C11** (Scheme C.3.1).

Reduction followed by condensation with the tripropyl derivative of Kemp's triacid imide acid chloride²² gave the autocatalytic precursor **C13**. Hydrolysis of the methyl ester proved to be problematic, due to both reduced reactivity and solubility of the biphenyl substituent. Acid **C14** was produced in quantitative yield by S_N2 dealkylation⁸⁷ of the methyl ester with thiolate anion. Formation of the acid chloride was unsuccessful with thionyl chloride due to destruction of the imide group, but proceeded at room temperature using oxalyl chloride. Condensation with 5'-amino-5'-deoxy-2',3'-isopropylideneadenosine produced the self-complementary reaction template **C15**.

The biphenylcarbazole-diimide structure provides very high affinity for 2',3'-isopropylideneadenosine. As the large size of the molecule leads to highly broadened 1H NMR spectra during titration, the binding constant in $CDCl_3$ can only be estimated as $10^5 M^{-1}$. The system is readily soluble in tetrahydrofuran and the association constant between **C13** and 2',3'-isopropylideneadenosine was measured by 1H NMR titration in $THF-d_8$ to be $576 M^{-1}$ at 298 K, based on the downfield shift of the imide proton from 9.65 ppm to 12.87 ppm. Titration in methanol was not possible due to limited solubility of ester **C13**.

With an association constant on the order of $10^5 M^{-1}$ in chloroform, the dimerization constant could be expected to on the order of $10^{10} M^{-1}$. Variation of the temperature of a 5 mM $CDCl_3$ solution of autocatalytic template **C15** shows two sharp imide peaks in the 1H NMR spectrum centered at 13 ppm at $-55^\circ C$. The imide resonances begin to broaden at $-25^\circ C$, with greater broadening in the downfield imide (predicted to be the Hoogsteen-bound imide based on the structural studies of Section A.2.2), and



Scheme C.3.2. Synthesis of tripropyl biphenylscorpion template.

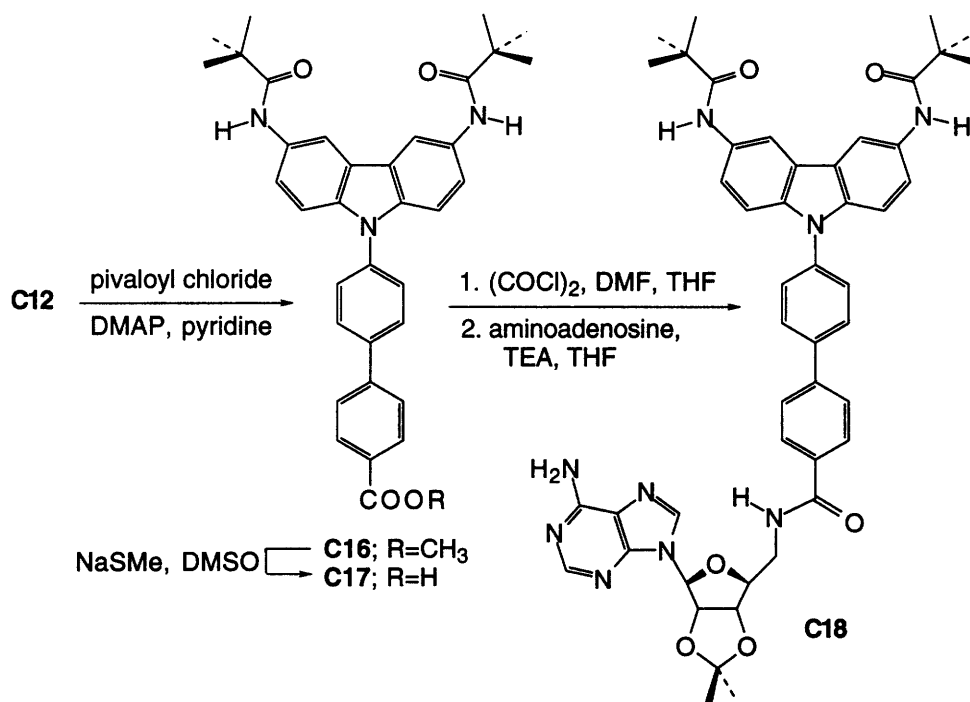
coalesce to a single broad resonance between 35 °C and 45 °C (Figure E.11b). With a maximum $\Delta\nu$ of 311 Hz (-55 °C), the barrier to conversion between the two imides at the coalescence temperature (taken to be 40 °C) is approximately 14.3 kcal/mol (using equations A.2.1 and A.2.2). Since no other binding modes are possible, this behavior can be considered the paradigm for a purely dimeric system. Using **C13** as a reference, vapor phase osmometry indicates a molecular weight in solution of 4205 amu, 163% of the value expected for a dimeric species (2585.1 amu). Again, the VPO measurement overestimates the weight of a polar species due to nonspecific aggregation.

In THF- d_8 , the dimerization constant would be expected to be in the vicinity of $3 \times 10^5 \text{ M}^{-1}$. While the ^1H NMR spectrum of **C15** in THF- d_8 does exhibit concentration dependence from 0.057-5.081 mM,

broadening of the resonances makes it impossible to calculate a dimerization constant.

The dimerization of **C15** appears so strong, in fact, that dimerization occurs even in DMSO, despite the highly competitive nature of this solvent. A ^1H NMR dilution study of the molecule in DMSO-d_6 from 0.068 mM to 6.13 mM was used to calculate a dimerization constant of approximately 169 M^{-1} . Hydration of DMSO makes it difficult to perform this at conditions of constant water content. Small variations of the water content (0.14-0.42%, based on relative integration of water and solvent peaks) have a significant effect on the shift of the imide proton. Comparison of the resonances of **C15** from 0.068 mM to 19.41 mM in DMSO-d_6 , however, show the following chemical shift changes: *i*) downfield shift of the imide proton, *ii*) upfield shifts of the adenine aromatic protons, *iii*) downfield shift of carbazole C^4H , *iv*) downfield shifts of the adenine NH_2 protons, and *v*) upfield shifts of the ribose 1' and 2' protons. Though the magnitude of these changes is very small (~ 0.01 ppm), these chemical shifts are typical of adenine binding within the diimide pocket, on top of the carbazole and biphenyl surfaces. This is the first observation of non-ionic hydrogen-bonding in DMSO seen in this laboratory.

The 'declawed' biphenylscorpion **C16** was used to synthesize 'template' **C18** (Scheme C.3.3) as a control for the role of imide binding in autocatalysis.



Scheme C.3.3. Synthesis of control for the biphenylscorpion autocatalytic system.

With all of the components in hand, kinetic examination of Reaction I (Figure C.3.2) ensued. All kinetic studies were performed by Edward Wintner. These are still in progress.

The reaction is autocatalytic; addition of the product, **C15**, results in an increase in the reaction rate.

Furthermore, this rate enhancement is solvent-dependent. Figure C.3.3 shows the rate enhancements produced by addition of the product (0.5 equiv) to the reaction between 5'-amino-5'-deoxy-2',3'-isopropylideneadenosine and the dinitrophenyl ester of **C14**. This system was examined under dilute conditions (0.07 mM ester, 0.4 mM amine, 8 mM TEA) by monitoring the release of dinitrophenol using UV/Vis spectroscopy. When there is no THF in the reaction mixture, the rate enhancement resulting from the addition of 0.5 equiv of template is reasonably small. The rate enhancement increases as the proportion of THF in the solvent increases, but no rate enhancement *at all* is seen above 50% THF; that is, there is no autocatalysis in these solvent conditions. This behavior can be assigned to a combination of the strong dimerization of the template **C15** and rate acceleration by polar solvent. In pure chloroform, there would be little free template due to the very large association constant. As the percentage of THF increases, dissociation of the dimer becomes less unfavored (the association constant between the two starting materials is reduced to 576 M^{-1} in pure THF). However, simultaneously, the background reaction rate increases with increasing solvent polarity. If the autocatalytic rate enhancement is small, then its effect may get 'lost' as the rate of the background reaction increases.

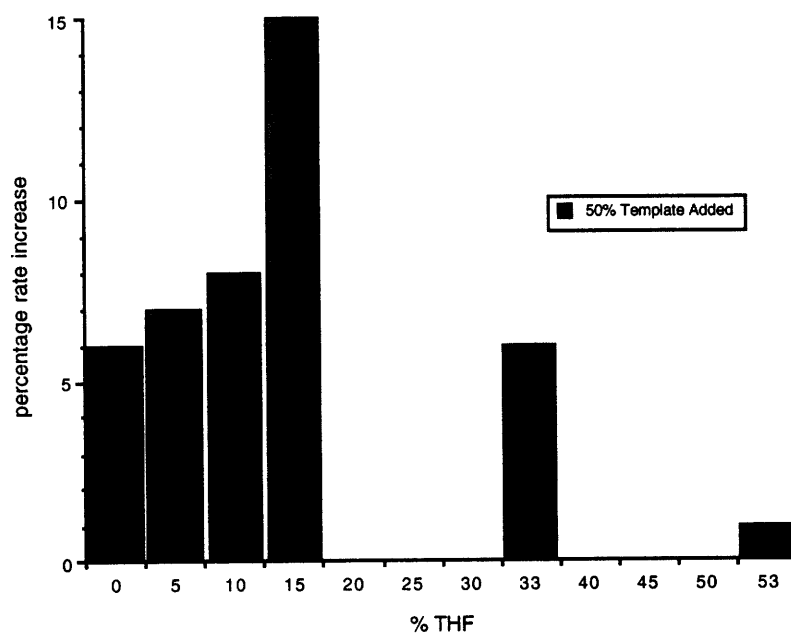


Figure C.3.3. Catalytic enhancement of autocatalysis in chloroform/THF mixtures.

Assuming that the rate enhancement is due to a template effect, the predicted structure of the tetrahedral intermediate (Figure C.3.4) was generated by minimization of the neutral tautomer within MacroModel⁷⁰ using the Polak-Ribiere¹²⁷ conjugate gradient algorithm and the AMBER* force field to a low energy gradient ($8.91 \times 10^{-4} \text{ kJ/mol-Å}$). It can be seen that the tetrahedral intermediate itself is in a relatively

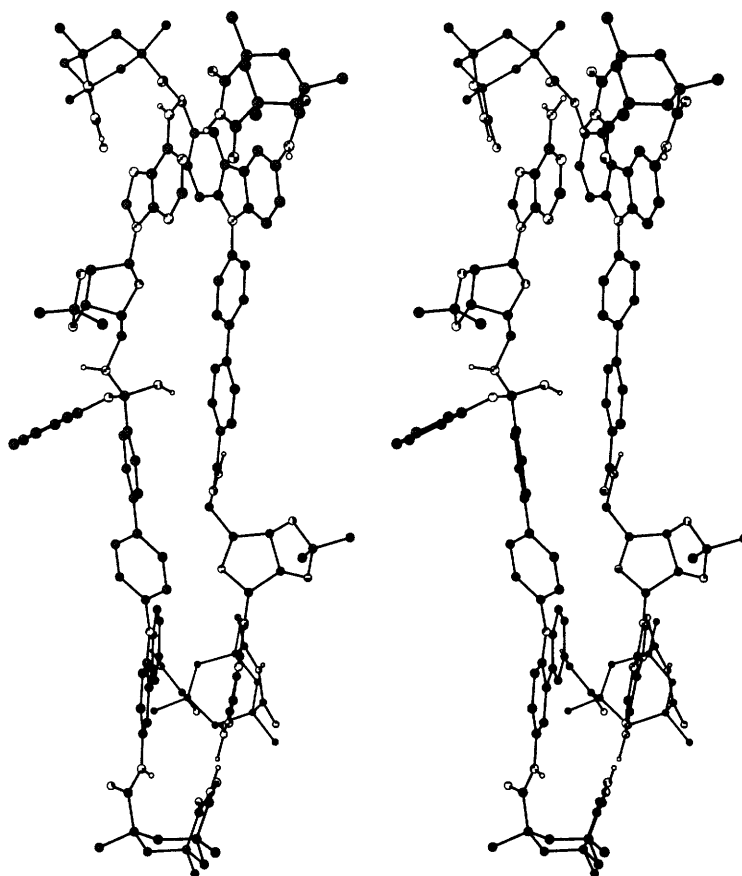


Figure C.3.4. Predicted structure of the autocatalytic tetrahedral intermediate (stereoview).

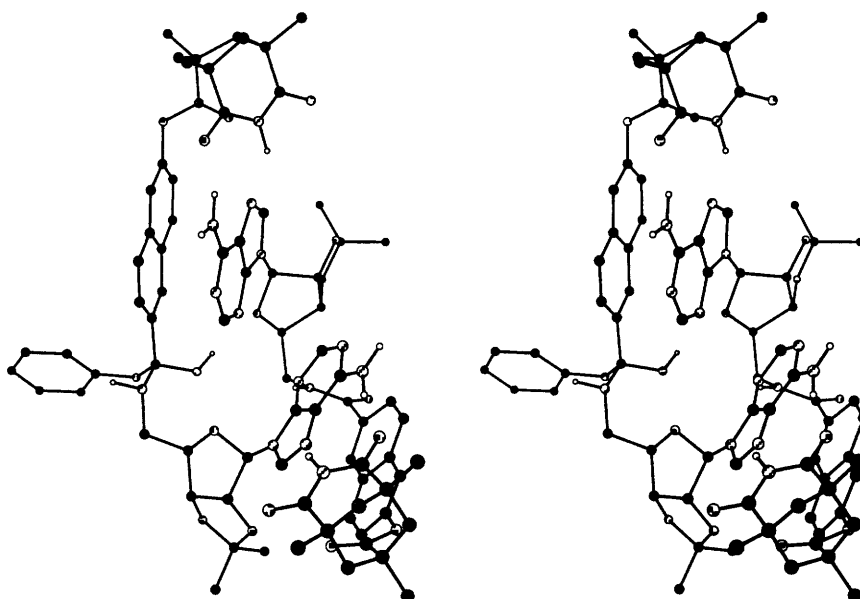


Figure C.3.5. Predicted structure of a tetrahedral intermediate in monoimide autocatalysis (stereoview).

hydrophobic environment, directly above the distal phenyl ring of the template and far removed from any polar functional groups. This is in sharp contrast to the predicted tetrahedral intermediate structure of

the naphthyl-monoimide self-replicating system (Figure C.3.5).^{108,116} If the breakdown of the tetrahedral intermediate is the rate-limiting step,¹¹¹ then a polar environment that facilitates proton transfer in the tetrahedral intermediate would be advantageous. The hydrophobic environment of the termolecular complex would seem to detract from the overall catalysis generated by the entropic contribution of a template effect. The termolecular reaction rate would likely be less solvent-dependent than the background reaction so that with an increase in the polarity of the medium, reaction within the termolecular complex would become less important to the overall rate than the background reaction.

Detailed studies of the autocatalytic reaction were performed by monitoring the initial rate of product formation by normal-phase analytical HPLC for the first 15-20% of reaction. The rates determined by linear regression of peak area versus time are reproducible within 5%. Within a single experiment, the correlation coefficient of the linear curve-fit is greater than 0.99. The concentrations of 5'-amino-5'-deoxy-2',3'-isopropylideneadenosine and the pentafluorophenyl ester of **C14** were 6.2 mM in 13% THF/CHCl₃. The triethylamine concentration was 96 mM. The template, **C15**, behaved similarly by variable-temperature ¹H NMR in 13% THF-d₈/CDCl₃ as in pure CDCl₃, indicating that the addition of THF does not significantly perturb the self-association of this molecule. Assuming that the effect of THF on binding is linear from 0% (10⁵ M⁻¹) to 100% THF (576 M⁻¹), then the association constant between diimide **C14** and adenosine can be expected to be approximately 87000 M⁻¹ in 13% THF/CHCl₃.

added 'catalyst' (0.5 equiv)	rate (10 ⁻³ mM/min)	relative rate
nothing	1.695	1.00
C15	2.617	1.54
C13	1.468	0.87
C18	8.315	4.91

Table C.3.1. Reaction rate of aminoadenosine with pentafluorophenyl ester of **C14** in 13% THF/CHCl₃.

The results of this study are shown in Figure C.3.6 and, in more detail, in Table C.3.1. Addition of the reaction product **C15** (0.5 equiv) produces a 54% increase in the reaction rate between the pentafluorophenyl ester of **C14** and 5'-amino-5'-deoxy-2',3'-isopropylideneadenosine. This is evidence for the autocatalytic nature of the reaction.

A number of control studies were performed to identify the mechanism by which autocatalysis operates. The addition of methyl ester **C13** produces a marginal reduction in the reaction rate, due to binding of free aminoadenosine to form an unproductive complex. There is only a minor effect from this addition as the high binding affinity of the diimide cleft ensures that the aminoadenosine is *already* bound by the active ester cleft. The distance between amino group and ester in the complex prevents these complexes from providing a pathway for the formation of product. That is, there is no preassociative reaction pathway.

Addition of valerolactam surprisingly produces no significant increase in reaction rate, unlike in the

monoimide system discussed in Section C.1. It may well be that the presence of THF in the medium accelerates the reaction sufficiently such that the contribution of valerolactam is minor. Certainly the lack of effect by a known aminoacylation catalyst suggests that the *trans*-amides are not effective as catalysts in this system.

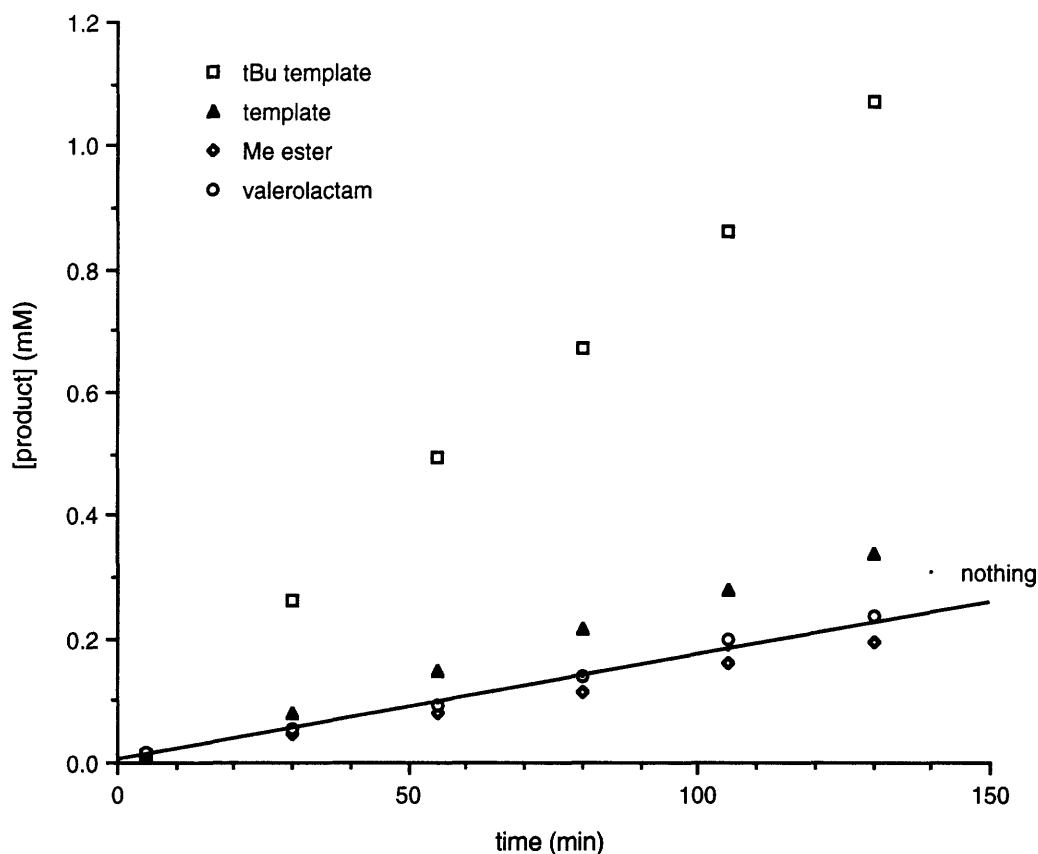


Figure C.3.6. Reaction rate of autocatalysis in 13% THF/chloroform.

It should be noted that in the normal-phase column used for these HPLC studies, all acylated-aminoadenosine compounds have approximately the same retention time. This means that in experiments where a combination of **C15** and **C18** are present, the *combined* peak of the two compounds was followed. There is the concern that this complicates the assay so that effects outside catalysis are observed. UV spectra were measured (Figure E.12) for solutions of **C15** and **C18** (10 μmol in 4% MeOH/ CHCl_3) to ensure no significant change in absorbance was responsible for the large catalytic effects of these molecules. A combined spectrum of the two compounds was not at all significantly different from the mathematical sum of the spectra of the two compounds alone.

The 'declawed' template **C18** was added (0.5 equiv) as a control to ascertain that recognition by the diimide cleft is a requirement for autocatalysis. This control accelerates the reaction significantly

better, *three*-fold, than the actual product. This goes against previous results obtained for the monoimide systems which show that neither the adenine nucleus nor amide groups in general catalyze aminoacylation. If this molecule were acting only by releasing aminoadenosine from the diimide cleft by its own complexation, then the reaction rate could not be more than the intrinsic background rate (see below, Table C.3.2). In fact, this reaction rate is more than twice the intrinsic background.

The chemical behavior of the *t*-butylanilide group is unusual and suggests that there are interactions occurring with the derivatives of **C15** that do not occur with the cleft. There is partial decomposition of **C18** to a second product on standing over a period of weeks. This decomposition is not evident for any of the carbazole-diimide clefts. Furthermore, the decomposition is independent of the adenosine end as the pentafluorophenyl ester of **C17** also decomposes and gives the same HPLC trace as **C18** after reaction with 5'-amino-5'-deoxy-2',3'-isopropylideneadenosine. Because of this behavior, any results obtained with the 'declawed' scorpion **C18** must be discounted.

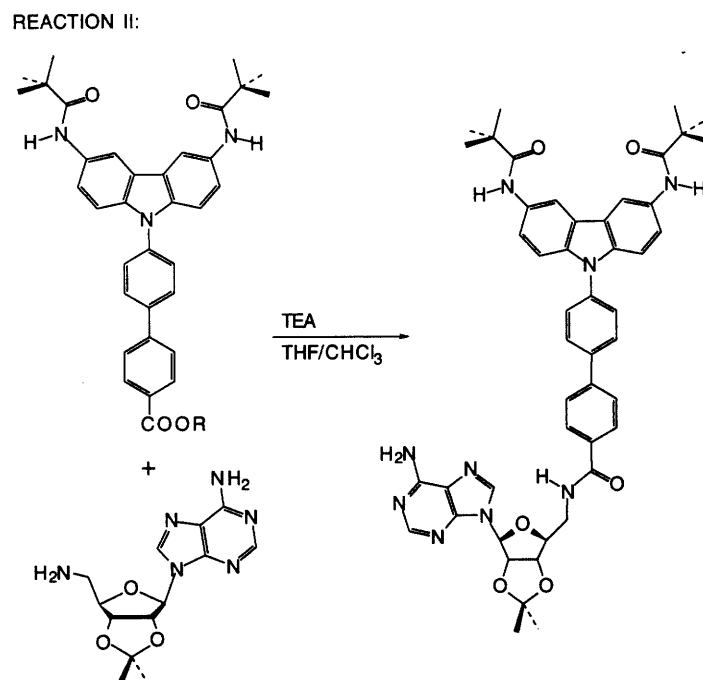


Figure C.3.7. The intrinsic background reaction for autocatalysis.

Due to the complication of high binding affinity and restricted access in these large complexes, kinetic studies were performed on the background reaction, Reaction II (Figure C.3.7). The aminoacylation of 5'-amino-5'-deoxy-2',3'-isopropylideneadenosine with the pentafluorophenyl ester of **C17** simply shows the intrinsic chemical reactivity of the active ester, without any complexation effects. This data is shown in Figure C.3.8 and, in detail, Table C.3.2. This reaction is over twice as fast as the reaction of the active ester of **C14**. Complexation of aminoadenosine by the diimide cleft blocks approach to the nucleophile from one side, thereby reducing access by the electrophile. This effect is reproduced by addition of **C13**

(1 equiv) to the background reaction (Table C.3.2) where one equivalent of the diimide cleft binds the aminoadenosine and causes a 44% reduction in reaction rate. The rate of the autocatalytic reaction itself (Reaction I) is slower yet possibly because the active ester also exists in solution as a complex with both hindered approach and a large molecular weight (~1500 amu). Addition of the reaction product **C18** (0.5 equiv) causes a 17% improvement in the reaction rate relative to the diimide-inhibited rate, but not over the background reaction. Since **C18** is not much more than a large acylated-aminoadenosine, this can be best assigned simply to a re-proportioning of adenines bound within the diimide cleft. The addition of **C18** with **C13** (1 equiv) causes the amount of free, unencumbered, aminoadenosine in solution to increase, thereby increasing the rate of coupling.

additive	rate (10 ⁻³ mM/min)	relative rate
nothing	3.350	1
C13 (1 equiv)	2.196	0.66
C13 (1 equiv) C18 (0.5 equiv)	2.564	0.77
C13 (1 equiv) C15 (0.5 equiv)	4.971	1.484

Table C.3.2. Reaction rates of aminoadenosine with pentafluorophenyl ester of **C17** in 13% THF/CHCl₃.

Finally, the autocatalytic template **C15** (0.5 equiv) was added with **C13** (1 equiv) to ascertain whether the net effect of this template is to simply to increase the concentration of free aminoadenosine. In fact, the reaction is 1.5 times the rate of the intrinsic background reaction and almost twice as fast as the reaction with **C18**. In these conditions, **C15** should exist solely as the dimer. Even so, any monomer present would be expected to have a net null effect, as the template contains both adenine and diimide-binding subunits. Furthermore, even if we assume that **C15** acts on Reaction II only by liberating free aminoadenosine (the behavior ascribed to the catalytic effect of **C18**), the rate of reaction *cannot* exceed the intrinsic background reaction. However, it is not unreasonable to expect that a complex that corresponds, in size, to a peptide of ~22 amino acids would have some effect on a general aminoacylation reaction. Moreover, knowing that the chemistry of the *t*-butylanilide is unusual when combined with the carbazole-diimide, little relevance can be assigned to this result.

We seem to have here a situation where the product from one reaction is a significantly better catalyst for the other reaction rather than its own. This could be the basis for the creation of an artificial hypercycle,¹²⁸ provided that molecular recognition could be shown as the origin of catalysis. The nature of the acatalytic effect occurring between these two systems is currently unknown.

The role of the individual functional groups of template **C15** in the rate acceleration by the template was probed by investigating the effect of small, non-complexing additives on Reaction I. Imide **VII** provides a test of the involvement of the imide group and the Kemp's anilide in autocatalysis. The affinity of the monoimide for adenine is *much* lower than that of the carbazole-diimide, so it can be expected that the

VII exists essentially free in the reaction solution. Amide **IX** (Figure C.3.9) presents the identical steric and electronic environment as the central amide of **C15**, but avoids the involvement of complexation and decomplexation effects. Litvinenko has shown that amides can act as catalysts in aminoacylation reactions of acyl chlorides in benzene through O-nucleophilic attack by the amide carbonyl at the reactive center.¹¹⁰ Amide **IX**, then, provides a direct test of the catalytic effect of amides in the autocatalytic system. As all results with 'declawed' template **C18** are suspect, 9-ethyladenine was added to the reaction to investigate the extent that complexation effects slow the initial reaction.

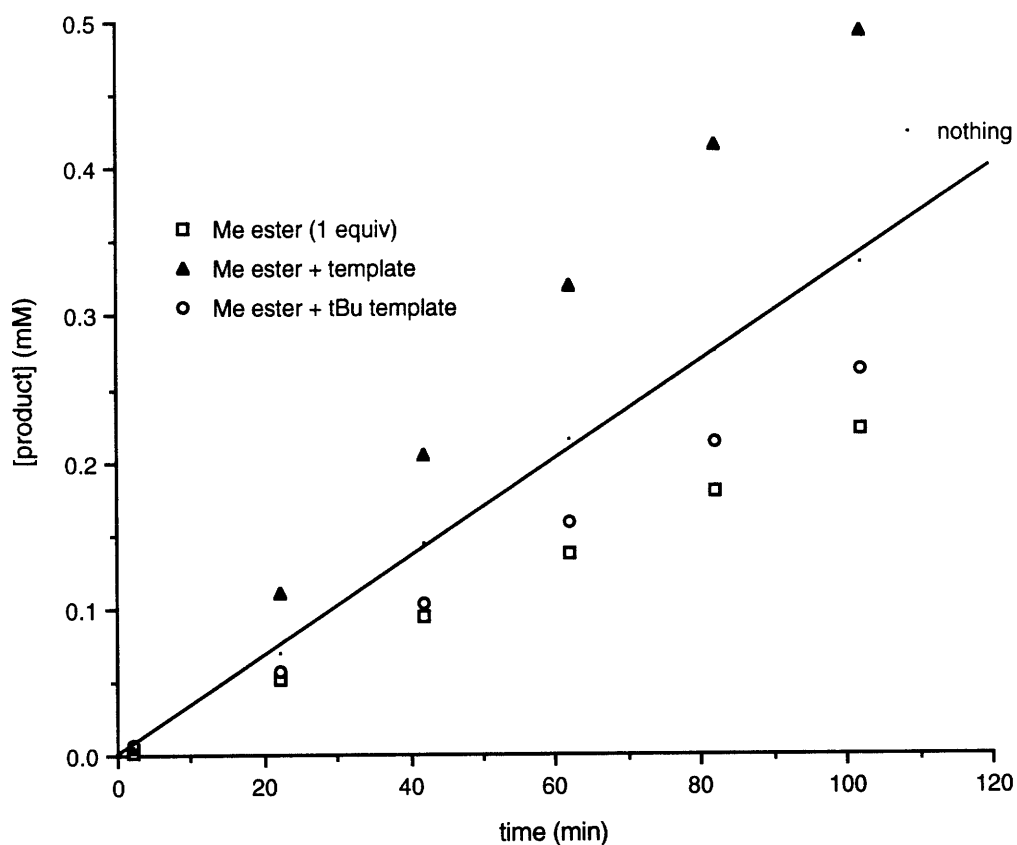


Figure C.3.8. Control studies with the pentafluorophenyl ester of **C17**.

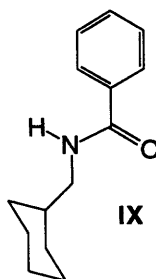


Figure C.3.9. Additional control for autocatalysis.

The results of this study are summarized in Figure C.3.10 and, in detail, in Table C.3.3. Imide **VII** has no effect at all on the reaction rate. This demonstrates that neither the imide, nor the Kemp's anilide are chemical catalysts of aminoacylations under these conditions. Furthermore, amide **IX** exerts no catalytic effect on the reaction and is even slightly inhibitory. While the origin of this inhibition is unknown, this demonstrates that a secondary amide in the same context as that in the template is *not* an aminoacylation catalyst.

added 'catalyst'	rate (10^{-3} mM/min)	relative rate
nothing	1.722	1.00
VII (0.5 equiv)	1.728	1.00
IX (1 equiv)	1.511	0.88
9-ethyladenine (1 equiv)	2.263	1.31

Table C.3.3. Control reactions of aminoadenosine with pentafluorophenyl ester of **C14** in 13% THF/ CHCl_3 .

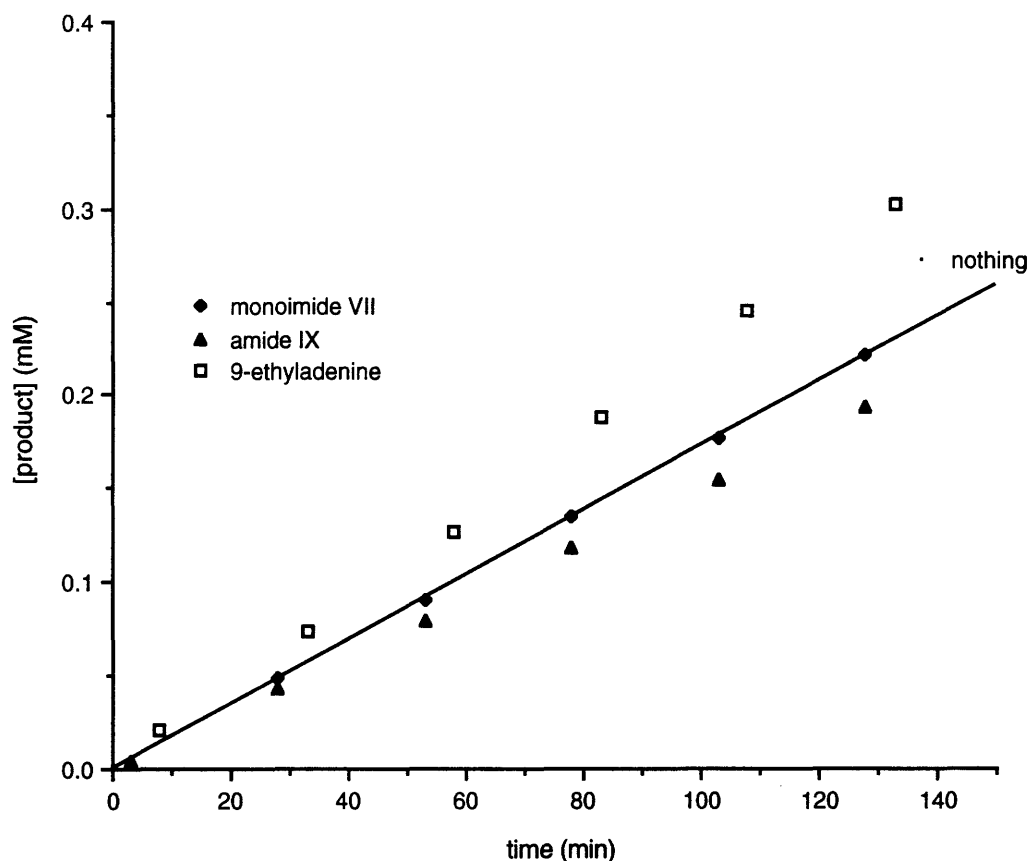


Figure C.3.10. The structural elements of the template are not catalysts, individually.

There was a small rate enhancement when 9-ethyladenine was added to the reaction. This is due to the competition of 9-ethyladenine for the diimide sites originally occupied by aminoadenosine. The decomplexed aminoadenosine reacts more rapidly with the active ester when free than when bound. The

extent of acceleration, however, is significantly lower than seen when **C18** was added. Once again, this highlights the unusual nature of the chemistry of the 'declawed' scorpion moiety. It should be noted that both amide **IX** and 9-ethyladenine were added to the reaction as a full equivalent. Under the standard conditions (0.5 equiv added), the effect of these additives could be expected to be at least half as great.

When taken together, the results of this study indicate that the individual components of template **C18** are not capable of catalyzing the reaction. In other words, the whole is greater than the sum of its parts. This provides a strong case for autocatalysis through the involvement of molecular recognition and a termolecular complex.

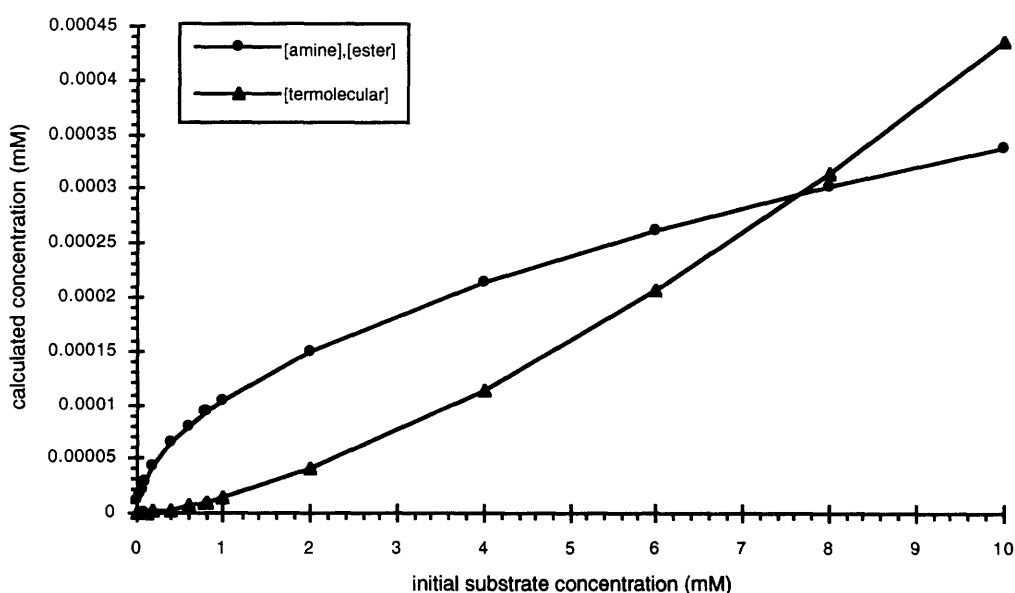


Figure C.3.11. Variation of free substrate and termolecular complex concentration with starting concentration.

There is an important consideration that arises from the high binding affinity of these diimide clefts when applied to autocatalysis. In the monoimide self-replicating systems, the dimerization of the template is of the same order of magnitude as the single binding event (adenine to imide). This allows the formation of the termolecular complex to compete with self-dimerization of the template. In the case of the diimide cleft, where the dimerization seems extraordinarily strong, then the concentration of the termolecular complex will be extremely small. In the solvent system used in these studies, the single binding event can be estimated to be 87000 M^{-1} and the dimerization of the template taken to be the square of that value. By inserting these values into a model of autocatalysis developed by James Nowick,¹¹⁶ the predicted concentration of the termolecular complex can be calculated as a function of initial substrate concentration with 0.5 equiv of template added (Figure C.3.11). Under the conditions of this study (6.2 mM substrate, 3.1 mM template), the concentration of the termolecular complex is predicted to be only $0.2 \text{ } \mu\text{mol}$. This is on the same order as the free (*i.e.* unbound) amine and ester. The reaction rate within

the termolecular complex would have to be significantly enhanced for the autocatalytic to be apparent if the termolecular complex is responsible for the rate enhancements seen.

Given that *self-replication* can be demonstrated as responsible for the autocatalysis observed here, improvements to this system would address the issue of small concentration of the termolecular complex by either moderating the dimerization of the template to levels more similar to the single binding event, or by improving reaction within the termolecular complex sufficiently for its effect on rate to be more important. Modifications along this line would require the placement of catalytic groups (e.g. *cis*- or primary amides) on the distal phenyl ring in a position to interact with the tetrahedral intermediate as it forms in the termolecular complex.

4. WATER-SOLUBILIZATION OF AN AUTOCATALYTIC SYSTEM

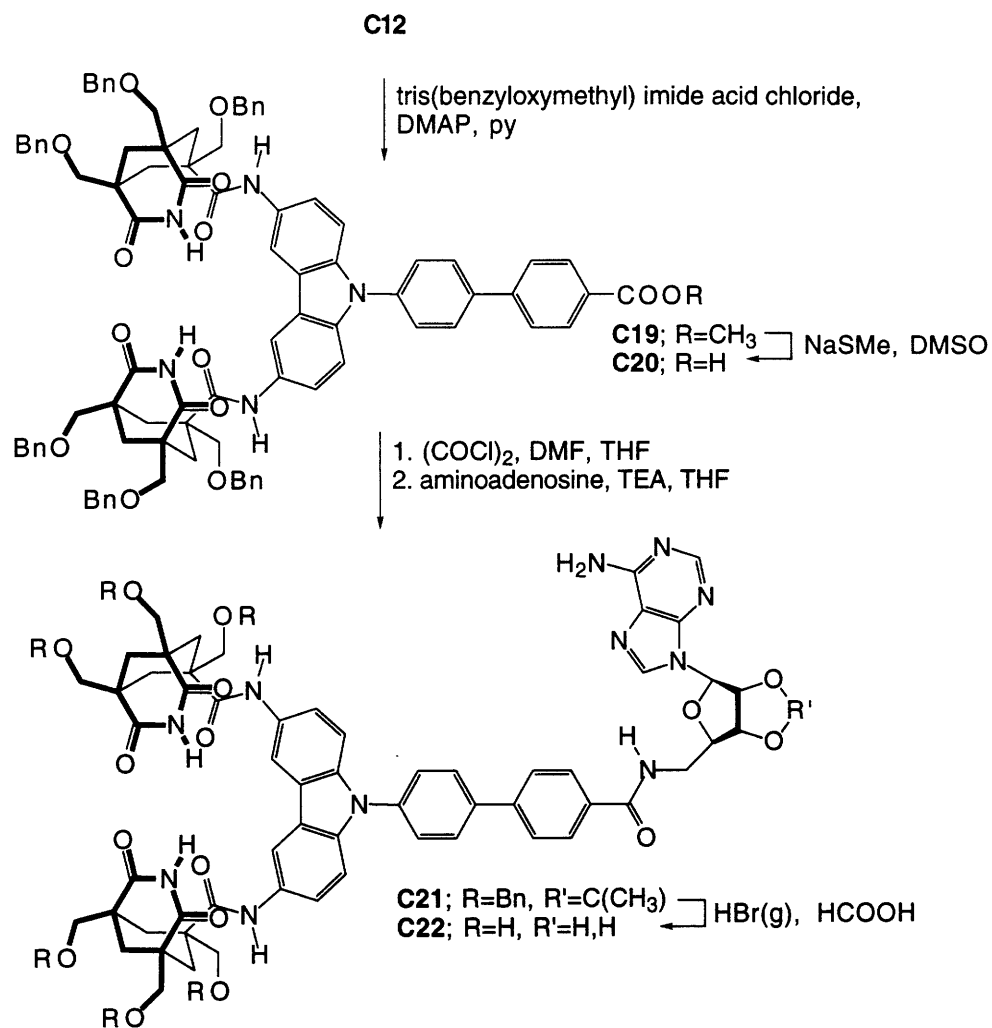
We attempted to extend this autocatalytic design to aqueous solution, both to alleviate complications from excessively high affinity and to improve the relevance of the system to biological comparison. For this, we exploit the tris((benzyloxy)methyl) derivative of Kemp's triacid,⁴² which can be deprotected by brief treatment with HBr(g) to cleanly give the water-soluble tris(hydroxymethyl) derivative. Though the strength of hydrogen bonding is minimal in water,⁴² the carbazole-diimide is capable of binding adenosine derivatives in aqueous solution. We have shown that a water-soluble carbazole-diimide complexes with adenosine with a binding affinity of 2.8 kcal/mol at 10 °C in water.⁴³ Admittedly, though, the major driving force for binding comes from hydrophobic, rather than hydrogen-bonding, forces.

Tris((benzyloxy)methyl) Kemp's triacid imide acid chloride was condensed with the 3,6-diaminocarbazole **C12** to give the protected cleft **C19**. Dealkylation of the ester, followed by coupling with amino-adenosine, gave the protected version of the autocatalytic template, **C21**. Brief treatment of this with HBr(g), gives the water-soluble reaction template **C22**.

Due both to solubility limitations and competition from hydrolysis, activated phenyl esters are not appropriate for autocatalysis in aqueous solvents. Accordingly, the water-soluble acid **C23** (Scheme C.3.5) was used in conjunction with the water-soluble coupling reagent, EDC (1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride) for condensation with 5'-amino-5'-deoxyadenosine.

A significant complication is the competition from attack by water and primary and secondary hydroxyl groups on the activated acid. Acid is regenerated when the nucleophile is water, but hydroxyl attack will result in the formation of non-self-complementary ester by-products. This competition significantly complicates the analysis of these systems. A trial reaction between acid **C23** and 5'-amino-5'-deoxyadenosine with EDC in DMF led to a complex mixture of products. Preparative reverse-phase HPLC of this mixture gave only a low yield of template **C22**. However, it was hoped that, in aqueous conditions, the small amount of template produced would be active enough as a catalyst to preferentially accelerate

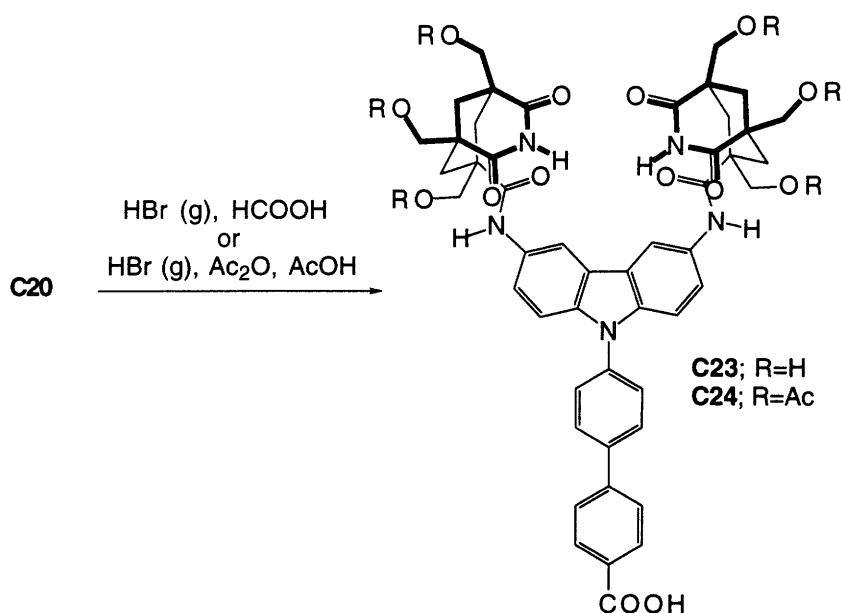
the target reaction over nonspecific processes.



Scheme C.4.1. Synthesis of tris(hydroxymethyl) biphenylscorpion template

Analysis of this reaction was intractable. The low solubility of the biphenyl carbazole-diimide derivatives mandated both high pH and dilute concentration (0.5 mM in pH 9.5 CHES buffer). These conditions make NMR analysis of this system impossible, due to limits imposed by sensitivity and the base-catalyzed exchange of imide and amide protons with bulk water.¹²⁹ HPLC analysis of the system proved problematic because of aggregation on the column that precluded resolution of the reaction mixture.

En route to the water-soluble clefts, a tris(acetoxymethyl) derivative of Kemp's triacid imide was synthesized with the goal of generating a precursor that could be made water-soluble under mild conditions. The hexakis(acetoxy) cleft **C24** was isolated in less than 30% yield from the mixture of partially acetylated products resulting from HBr(g) treatment of the benzyl-protected acid **C20** in the presence of acetic anhydride. This mixture profile was reproduced by treatment of hydroxy acid **C23** with acetyl chloride. Unfortunately, the low yield of this conversion restricted the application of this approach.



Scheme C.4.2. Masked deprotection of tris((benzyloxy)methyl) biphenylscorpion.

5. EXPERIMENTAL

5.1 Synthesis

General

All commercially available compounds (Aldrich) were used without further purification unless otherwise indicated. DMSO- d_6 (99.9%-D, MSD Isotopes or Cambridge Isotope Labs) were used from freshly opened vials. CDCl_3 (99.8%-D, MSD or CIL) was stored over molecular sieves and passed through dry basic alumina just prior to use. Tetrahydrofuran used in anhydrous conditions were distilled from sodium-benzophenone ketyl; dimethyl sulfoxide was distilled from sodium hydroxide and stored over molecular sieves (4 Å). Thin layer chromatography was performed on Merck Silica 60 F-254 precoated TLC plates. Column chromatography was performed on Merck Silica Gel 60 (230-400 mesh) according to Still *et al.*⁵⁹ Glassware used for anhydrous conditions either was baked overnight at 150 °C, assembled hot, cooled under vacuum, and filled with argon; or was flamed dry under vacuum, cooled and filled with argon before use. Standard inert-atmosphere techniques were used for solvent transfers by syringe.

^1H NMR spectra were recorded on Varian XL-300, UN-300, and VXR-500 spectrometers in solvents as indicated. Chemical shifts are reported as parts per million (δ) relative to residual solvent protons. Melting points were obtained on an Electrothermal IA9100 digital melting point apparatus and are calibrated. IR spectra were recorded on a Perkin-Elmer series 1600 FT-IR spectrometer. High resolution mass spectra were obtained on a Finnegan MAT 8200 mass spectrometer.

Adenosyl tripropyl methylscorpion (C1)

N-(Carboxymethyl)-3,6-bis[*(cis,cis*-2,4-dioxo-1,5,7-tripropyl-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino] carbazole **A6** (515.5 mg, 0.595 mmol) was stirred under Ar with CDI (1.09 equiv) in anhydrous THF for 3.3 h. 5'-Amino-5'-deoxy-2',3'-(isopropylidene)adenosine¹³⁰ (1.15 equiv) was added and the solution was stirred at room temperature for 44 h. Solvent was removed by rotary evaporation and the residue purified by flash chromatography (20-30% MeOH/CHCl₃) to yield 237.5 mg of pure cleft, after slow evaporation from CH₂Cl₂/hexanes. The mixed fractions were concentrated and triturated in MeOH to yield 76.6 mg of pure cleft, after filtration and drying. The combined yield was 314.1 mg (46%): IR (KBr) 3377, 3214, 2959, 1700, 1696, 1647, 1533, 1468, 1201 cm⁻¹; ¹H NMR 300 MHz, DMSO-d₆) δ 10.685 (s, 2 H, imide), 9.118 (s, 2 H, amide), 8.306 (s, 2 H, amine), 8.130 (d, 2 H, J = 1.6 Hz), 8.051 (s, 1 H, amide), 7.39 (m, 4 H), 7.302 (s, 1 H), 7.272 (s, 1 H), 6.150 (d, 1 H, J = 2.9 Hz), 5.42 (m, 1 H), 4.946 (s, 2 H), 4.91 (m, 1 H), 4.21 (m, 1 H), 3.4 (m, 2 H), 2.680 (d, 4 H, J = 13.5 Hz), 2.041 (d, 2 H, J = 12.7 Hz), 1.8 (m, 4 H), 1.536 (s, 3 H), 1.51 (m, 4 H), 1.4-1 (m, 22 H), 1.309 (s, 3 H), 0.877 (t, 12 H, J = 6.9 Hz), 0.802 (t, 6 H, J = 7.2 Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₆₃H₈₄N₁₁O₁₀ (M+H), 1154.6403; found, 1154.6408.

N⁶-Methyladenosyl tripropyl methylscorpion (C2)

The methylscorpion acid **A6** (196.9 mg, 0.2273 mmol) was condensed with N⁶-methyl-5'-amino-5'-deoxy-2',3'-(isopropylidene)adenosine (1.015 equiv, prepared according to Kolb et al.¹³⁰ by J.-I. Hong) in the presence of EDC (1.028 equiv) in anhydrous THF (3 mL) for 21.5 h at room temperature under Ar. After removal of solvent under reduced pressure, the residue was purified twice by flash chromatography (1.5% MeOH/25% EtOAc/CH₂Cl₂ and 45% EtOAc/CH₂Cl₂) to give 115.6 mg (44%) of white powder **25**: IR (KBr) 3375, 2958, 1700, 1624, 1534, 1490, 1466, 1200 cm⁻¹; ¹H NMR (300 MHz, DMSO-d₆) δ 10.397 (s, 2 H, imide), 9.142 (s, 2 H, amide), 8.570 (s, 1 H, amide), 8.343 (s, 2 H, amine), 8.293 (s, 1 H), 8.103 (d, 2 H, J = 1.8 Hz), 7.839 (s, 1 H), 7.427 (dd, 2 H, J = 8.8, 1.6 Hz), 7.329 (d, 2 H, J = 8.4 Hz), 6.184 (d, 1 H, J = 2.7 Hz), 5.48 (m, 1 H), 4.94 (m, 1 H), 4.18 (m, 1 H), 3.4 (m, 2 H), 2.948 (m, 3 H), 2.666 (d, 4 H, J = 13.2 Hz), 2.025 (d, 2 H, J = 12.3 Hz), 1.785 (m, 4 H), 1.54 (s, 3 H), 1.51 (m, 4 H), 1.4-1 (m, 22 H), 1.31 (s, 3 H), 0.88 (t, 12 H, J = 7 Hz), 0.80 (t, 6 H, J = 7 Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₆₄H₈₆N₁₁O₁₀ (M+H), 1168.6559; found, 1168.6568.

Adenosyl tris((benzyloxy)methyl) methylscorpion (C3)

In an analogous fashion to **C1**, tris((benzyloxy)methyl) methylscorpion acid (102.2 mg, 0.07658 mmol, prepared by K. Kobiro by condensing N-(methoxy carbonylmethyl)-3,6-diaminocarbazole **A3** with tris((benzyloxy)methyl) Kemp's acyl chloride⁴²) was condensed with 1.21 equiv 5'-amino-5'-deoxy-2',3'-isopropylideneadenosine using 1.1 equiv CDI in 10 mL THF. After two flash columns (2% MeOH/CHCl₃ and 1% MeOH/CHCl₃), 62 mg (50%) of a beige powder were recovered which was used without further

purification: ^1H NMR (500 MHz, DMSO- d_6) δ 10.831 (s, 2 H, imide), 9.312 (s, 2 H, amide), 8.442 (s, 1 H), 8.337 (s, 2 H, amine), 8.129 (s, 1 H), 8.109 (s, 2 H), 7.4-7.2 (m, 34 H), 6.166 (d, 1 H, $J = 2.4$ Hz), 5.45 (m, 1 H), 4.967 (s, 2 H), 4.925 (m, 1 H), 4.53 (m, 12 H), 4.21 (m, 1 H), 3.596 (dd, 8 H, $J = 186, 8.8$ Hz), 3.444 (s, 4 H), 3.38 (m, 2 H), 2.568 (d, 4 H, $J = 14.2$ Hz), 2.342 (d, 2 H, $J = 12.2$ Hz), 1.6 (m, 6 H), 1.526 (s, 3 H), 1.300 (s, 3 H).

Adenosyl tris(hydroxymethyl) methylscorpion (C4)

The benzyl-protected cleft **C3** was dissolved in THF (2 mL) under Ar. Pd black (37 mL) and 6% HCOOH/MeOH (3 mL) were added.¹²² The suspension was stirred at room temperature for 45 h then filtered through Celite. Removal of solvent by rotary evaporation recovered 26.2 mg (97%) off-white product: IR (KBr) 3363, 2935, 2871, 1695, 1653, 1647, 1473, 1206 cm^{-1} ; ^1H NMR (500 MHz, DMSO- d_6) δ 10.676 (s, 2 H, imide), 9.131 (s, 2 H, amide), 8.318 (s, 2 H), 8.178 (s, 2 H), 8.082 (s, 1 H), 7.397 (m, 3 H), 7.283 (d, 2 H, $J = 8.5$ Hz), 6.160 (m, 1 H), 5.443 (m, 1 H), 5.10 (m, 1 H), 4.94 (m, 2 H), 4.77 (m, 2 H), 4.21 (m, 1 H), 3.779 (d, 3 H, $J = 9.5$ Hz), 2.406 (d, 4 H, $J = 13.5$ Hz), 2.207 (d, 2 H, $J = 13.0$ Hz), 1.541 (s, 3 H), 1.490 (dd, 4 H, $J = 13.5, 1.0$ Hz), 1.332 (d, 2 H, $J = 14.0$ Hz), 1.318 (s, 3 H); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for $\text{C}_{51}\text{H}_{60}\text{N}_{11}\text{O}_{16}$ (M+H), 1082.4220; found, 1082.4221.

Adenosyl tripropyl phenylscorpion (C5)

Phenyl scorpion acid **B6** (156 mg, 0.168 mmol) and 5'-amino-5'-deoxy-2',3'-isopropylidene-adenosine¹³⁰ (1.023 equiv) were stirred in anhydrous THF (4 mL) with EDC (1.043 equiv) for 20.5 h. Solvent was removed by rotary evaporation and the residue purified by flash chromatography (3% MeOH/ CHCl_3) to give 120.1 mg (58%) of cleft: IR (KBr) 3374, 2958, 1696, 1647, 1507, 1466, 1204, 1091 cm^{-1} ; ^1H NMR (300 MHz, DMSO- d_6) δ 10.345 (s, 2 H, imide), 9.218 (s, 2 H, amide), 8.806 (br t, 1 H, amide), 8.350 (s, 1 H), 8.288 (d, 2 H, $J = 1.8$ Hz), 8.114 (d, 2 H, $J = 8.4$ Hz), 8.106 (s, 1 H), 7.699 (d, 2 H, $J = 8.4$ Hz), 7.493 (dd, 2 H, $J = 9.0, 1.8$ Hz), 7.358 (d, 2 H, $J = 8.7$ Hz), 7.326 (s, 2 H, amine), 6.176 (d, 1 H, $J = 2.7$ Hz), 5.528 (dd, 1 H, $J = 6.4, 2.5$ Hz), 5.100 (dd, 1 H, $J = 6.3, 3.3$ Hz), 4.351 (m, 1 H), 3.585 (m, 2 H), 2.678 (d, 4 H, $J = 13.2$ Hz), 2.030 (d, 2 H, $J = 12.9$ Hz), 1.78 (m, 4 H), 1.546 (s, 3 H), 1.51 (m, 4 H), 1.33 (s, 3 H), 1.4-1 (m, 22 H), 0.875 (t, 12 H, 7.0 Hz), 0.805 (t, 6 H, 7.3 Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for $\text{C}_{68}\text{H}_{86}\text{N}_{11}\text{O}_{10}$ (M+H), 1216.6559; found, 1216.6554.

Trimethyl jackknife scorpion (C6)

EDC (1.19 equiv) was added to a solution of trimethyl methylscorpion acid (425.8 mg, 0.610 mmol, prepared by G. Deslongchamps) and 9-(2-aminoethyl)adenine¹³¹ (1.06 equiv) in anhydrous DMF (15 mL) under Ar. The reaction solution was stirred for 46 h at room temperature. Solvent was removed by rotary evaporation under reduced pressure and the residue purified by flash chromatography (5% MeOH/ CHCl_3) to give **C6** (215.4 mg, 41%) as an off-white solid: IR (KBr) 3393, 3204, 2970, 1693, 1654, 1521, 1465, 1315,

1207 cm⁻¹; ¹H NMR (300 MHz, DMSO-d₆) δ 11.155 (s, 2 H, imide), 9.187 (s, 2 H, amide), 8.125 (d, 2 H, J = 1.8 Hz), 8.002 (br t, 1 H, amide), 7.925 (s, 1 H), 7.879 (s, 1 H), 7.418 (dd, 2 H, J = 8.8, 1.9 Hz), 7.310 (s, 2 H, amine), 7.140 (d, 2 H, J = 9.0 Hz), 4.816 (s, 2 H), 4.179 (t, 2 H, J = 5.7 Hz), 3.598 (q, 2 H, J = 3.9 Hz), 2.774 (d, 4 H, J = 13.5 Hz), 1.942 (d, 2 H, J = 12.9 Hz), 1.436 (d, 2 H, J = 12.9 Hz), 1.263 (d, 4 H, J = 14.4 Hz), 1.218 (s, 6 H), 1.155 (s, 12 H); HRMS (EI) calcd for C₄₅H₅₁N₁₁O₇, 857.3973; found, 857.3968.

Tripropyl jackknife scorpion (C7)

Tripropyl methylscorpion acid **A6** (198.9 mg, 0.2296 mmol) and 9-(2-aminoethyl)adenine¹³¹ (1.097 equiv) were stirred at room temperature under Ar with EDC (1.04 equiv) in anhydrous THF (5 mL) for 22 h. Following removal of solvent by rotary evaporation, flash chromatography (0-1% MeOH/50% EtOAc/CH₂Cl₂) yielded 81.1 mg (34%) of pure **C7**: IR (KBr) 3382, 3213, 2930, 1695, 1653, 1528, 1467, 1307, 1208 cm⁻¹; ¹H NMR (300 MHz, DMSO-d₆) δ 11.132 (s, 2 H, imide), 9.131 (s, 2 H, amide), 8.108 (d, 2 H, J = 1.6 Hz), 7.982 (br s, 1 H, amide), 7.922 (s, 1 H), 7.894 (s, 1 H), 7.420 (dd, 2 H, J = 8.8, 1.8 Hz), 7.305 (s, 2 H, amine), 7.132 (d, 2 H, J = 8.8 Hz), 4.813 (s, 2 H), 4.172 (br t, 2 H), 4.169 (br s, 2 H), 2.700 (d, 4 H, J = 14.0 Hz), 2.054 (d, 2 H, J = 12.5 Hz), 1.79 (m, 4 H), 1.54 (m, 4 H), 1.4-1 (m, 22 H), 0.88 (t, 12 H, J = 7 Hz), 0.80 (t, 6 H, J = 7 Hz); HRMS (EI) calcd for C₅₇H₇₅N₁₁O₇, 1025.5851; found, 1025.5842.

4-Iodo-4-biphenylcarboxylic acid (C8)

Iodination of 4-biphenylcarboxylic acid was performed according to Merkushev and Yudina.¹³²

Methyl 4'-Iodo-4-biphenylcarboxylate (C9)

Thionyl chloride (50 mL) was added slowly dropwise over 2 h to a suspension of acid **C8** (14.49 g, 44.707 mmol) in dry MeOH under a drying tube. The suspension was refluxed for 4 h and stirred at room temperature for 24 h. Solvents were removed by rotary evaporation and then under vacuum. The crude was recrystallized from refluxing EtOAc (100 mL) and CHCl₃ (200 mL) to give 10.63 g (70%) of off-white platelets: ¹H NMR (500 MHz, CDCl₃) δ 8.102 (d, 2 H, J = 8.5 Hz), 7.796 (d, 2 H, J = 8.8 Hz), 7.617 (d, 2 H, J = 8.5 Hz), 7.356 (d, 2 H, J = 8.5 Hz), 3.943 (s, 3 H).

N-(4-(4'-(Methoxycarbonyl)phenyl)phenyl)carbazole (C10)

Iodobiphenyl **C9** (5.26 g, 31.43 mmol), carbazole (1.001 equiv), K₂CO₃ (1.001 equiv), and Cu powder (0.127 equiv) were intimately mixed with mortar and pestle. The solids were suspended in xylenes (20 mL) and mechanically stirred under reflux for 2 days. The reaction was slightly cooled, then briefly refluxed with benzene (250 mL). The suspension was cooled, stirred overnight, then filtered through Celite and reduced by rotary evaporation under reduced pressure. The residue was recrystallized from refluxing MeOH (50 mL) and EtOAc (200 mL) to give 5 g of off-white flaky crystals. The mother liquor was concentrated and again recrystallized from refluxing MeOH (30 mL) and EtOAc (60 mL) to give 1.11 g

product. The second mother liquor was concentrated and purified by chromatography (50% hexanes/CHCl₃) to give 269 mg of pure product. The combined yield of crystalline **C10** was 6.379 g (54%): mp 191.7-192.7 °C; IR (KBr) 2949, 1721, 1604, 1528, 1496, 1450, 1276, 1230, 1178, 1110, 835, 774, 753, 724 cm⁻¹; ¹H NMR (300 MHz, CDCl₃) δ 8.181 (d, 2 H, J = 8.6 Hz), 8.169 (d, 2 H, J = 7.5 Hz), 7.864 (d, 2 H, J = 8.6 Hz), 7.770 (d, 2 H, J = 8.6 Hz), 7.684 (d, 2 H, J = 8.8 Hz), 7.489 (d, 2 H, J = 7.7 Hz), 7.439 (td, 2 H, J = 7.5, 1.2 Hz), 7.315 (m, 2 H), 3.976 (s, 3 H); HRMS (EI) calcd for C₂₆H₁₉NO₂, 377.1416; found, 377.1416.

N-(4-(4-(Methoxycarbonyl)phenyl)phenyl)phenyl)-3,6-dinitrocarbazole (**C11**).

Carbazole **C10** (3.03 g, 8.03 mmol) was suspended in HOAc (60 mL) and brought to 70 °C in an oil bath. HNO₃ (conc., 30 mL) was added dropwise over 0.5 h under a drying tube. The yellow suspension was stirred at this temperature for 5 h and brought to room temperature for 3 h. The suspension was stirred with water (200 mL), filtered and dried to give 3.635 g (97%) yellow powder **C11** which was used without further purification: ¹H NMR (500 MHz, DMSO-d₆) δ 9.642 (d, 2 H, J = 2.5 Hz), 8.419 (dd, 2 H, J = 9.5, 2.5 Hz), 8.116 (d, 2 H, J = 8.5 Hz), 8.103 (d, 2 H, J = 8.0 Hz), 7.984 (d, 2 H, J = 8.5 Hz), 7.860 (d, 2 H, J = 8.5 Hz), 7.605 (d, 2 H, J = 9.5 Hz), 3.896 (s, 3 H).

N-(4-(4-(Methoxycarbonyl)phenyl)phenyl)phenyl)-3,6-diaminocarbazole (**C12**)

A suspension of dinitro **C11** (887 mg, 1.897 mmol) and 10% Pd-C (22 wt%) in THF (150 mL) was hydrogenated at balloon pressure for 26 h. After filtration through Celite, solvent was removed by rotary evaporation and the unstable yellow solid was carried on immediately: ¹H NMR (500 MHz, CDCl₃) δ 8.157 (d, 2 H, J = 8.5 Hz), 7.801 (d, 2 H, J = 9.0 Hz), 7.744 (d, 2 H, J = 8.5 Hz), 7.651 (d, 2 H, J = 9.0 Hz), 7.333 (d, 2 H, J = 8.5 Hz), 7.277 (d, 2 H, J = 2.5 Hz), 6.986 (s, 2 H, amine), 6.767 (dd, 2 H, J = 8.7, 2.2 Hz), 3.98 (s, 3 H).

N-(4-(4-(Methoxycarbonyl)phenyl)phenyl)phenyl)-3,6-bis[*[(cis,cis-2,4-dioxo-1,5,7-tripropyl-3-azabicyclo[3.3.1]non-7-yl)carbonyl]amino*]carbazole (**C13**).

Diamine **C12** and tripropyl Kemp's acyl chloride²² (2.05 mol equiv) were refluxed with catalytic DMAP (0.1 equiv) in pyridine (50 mL) for 16 h under Ar. Solvent was removed by rotary evaporation under reduced pressure and the residue taken up in EtOAc (200 mL) and CHCl₃ (25 mL). The combined organic layers were washed with 1 N HCl (40 mL), and brine (50 mL) and dried over Na₂SO₄. After concentration, the crude product was purified by gravity chromatography (2.5% MeOH/CHCl₃), followed by trituration in MeOH (20 mL). The yellowish solid was trituated a second time in MeOH (20 mL) to give 1.129 g (58%) of beige solid: IR (KBr) 3378, 3210, 2958, 1700, 1466, 1280, 1198, 1113, 774 cm⁻¹; ¹H NMR (500 MHz, DMSO-d₆) δ 10.368 (s, 2 H, imide), 9.232 (s, 2 H, amide), 8.312 (d, 2 H, J = 1.7 Hz), 8.096 (d, 2 H, J = 8.5 Hz), 8.048 (d, 2 H, J = 8.5 Hz), 7.966 (d, 2 H, J = 8.5 Hz), 7.738 (d, 2 H, J = 8.3 Hz), 7.494 (dd, 2 H, J =

9.0, 1.7 Hz), 7.379 (d, 2 H, J = 8.8 Hz), 3.892 (s, 3 H), 2.682 (d, 4 H, J = 13.9 Hz), 2.031 (d, 2 H, J = 12.7 Hz), 1.78 (m, 4 H), 1.51 (m, 4 H), 1.4-1.1 (m, 22 H), 0.876 (t, 12 H, J = 7.1 Hz), 0.807 (t, 6 H, J = 7.2 Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₆₂H₇₆N₅O₈ (M+H), 1018.5694; found, 1018.5690.

N-(4-(4'-Carboxyphenyl)phenyl)-3,6-bis(((cis,cis-2,4-dioxo-1,5,7-tripropyl-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino)carbazole (**C14**).

Anhydrous DMSO (10 mL) was added via syringe to ester **C13** (402 mg, 0.3948 mmol) and NaSMe (5.2 equiv) under Ar and the pale yellow solution was stirred at 80 °C for 1 h. After cooling to room temperature, the reaction was taken up in CHCl₃ (50 mL) and washed with 1 N HCl (2 x 30 mL), water (2 x 50 mL), and brine (50 mL), and dried over Na₂SO₄. Solvent was removed by rotary evaporation. The pale yellow solid was dried under vacuum to give 375 mg (95%) of acid **C14** which was used without further purification: ¹H NMR (500 MHz, DMSO-d₆) δ 10.365 (s, 2 H, imide), 9.230 (s, 2 H, amide), 8.312 (d, 2 H, J = 1.5 Hz), 8.072 (d, 2 H, J = 8.5 Hz), 8.036 (d, 2 H, J = 8.5 Hz), 7.931 (d, 2 H, J = 8.0 Hz), 7.730 (d, 2 H, J = 8.0 Hz), 7.491 (dd, 2 H, J = 8.5, 2.0 Hz), 7.377 (d, 2 H, J = 8.5 Hz), 2.682 (d, 4 H, J = 14.0 Hz), 2.030 (d, 2 H, J = 12.5 Hz), 1.80 (m, 4 H), 1.51 (m, 4 H), 1.35-1.1 (m, 22 H), 0.874 (t, 12 H, J = 7.0 Hz), 0.805 (t, 6 H, J = 7.2 Hz).

Activated phenyl esters of N-(4-(4'-carboxyphenyl)phenyl)-3,6-bis(((cis,cis-2,4-dioxo-1,5,7-tripropyl-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino)carbazole (typical conditions).

Anhydrous THF (~5 mL/100 mg acid) was added to a round-bottomed flask with acid **C14**, phenol (1.7-2.5 equiv), EDC (1.1-2.8 equiv), and DMAP (catalytic). The reaction mixture was stirred overnight at room temperature under Ar. Solvent was removed by rotary evaporation and the residue chromatographed on silica using EtOAc/CH₂Cl₂. Yields were typically 70-85%. Purity was estimated to be >95% by HPLC.

Tripropyl biphenylscorpion template (C15)

Acid **C14** (100 mg, 0.0996 mmol) was dissolved in anhydrous THF (7 mL). Oxalyl chloride (50 μL, 5.76 equiv) and catalytic DMF (1 μL) were added. The yellow solution was stirred under Ar at room temperature for 1 h and solvent was removed by rotary evaporation under reduced pressure. The acid chloride was pumped dry for 4 h. 5'-Amino-5'-deoxy-2',3'-isopropylideneadenosine¹³⁰ (1.18 equiv) was added to the reaction vessel and anhydrous THF (6 mL) was added to dissolve the reagents. Triethylamine (70 μL, 5.04 equiv) was added to the reaction solution accompanied by immediate formation of a white precipitate. After stirring at room temperature under Ar for 20.5 h, the reaction was filtered and solvents removed by rotary evaporation. Flash chromatography (4% MeOH/CHCl₃) gave partially pure product which was precipitated from an acetone solution (1 mL) with hexanes (5 mL) to give 78 mg (53%) of beige solid **C15**: IR (KBr) 3374, 3214, 2958, 1696, 1647, 1491, 1465, 1204, 1079 cm⁻¹; ¹H NMR (500 MHz, DMSO-d₆) δ 10.380 (s, 2 H, imide), 9.230 (s, 2 H, amide), 8.361 (s, 1 H), 8.318 (s, 1 H),

8.114 (d, 2 H, J = 1.5 Hz), 8.032 (d, 2 H, J = 8.5 Hz), 7.983 (d, 2 H, J = 8.5 Hz), 7.901 (d, 2 H, J = 8.5 Hz), 7.720 (d, 2 H, J = 8.0 Hz), 7.492 (dd, 2 H, J = 9.0, 1.5 Hz), 7.442 (br s, 2 H, amine), 7.373 (d, 2 H, J = 8.5 Hz), 6.174 (d, 1 H, J = 2.5 Hz), 5.507 (dd, 1 H, J = 6.0, 2.5 Hz), 5.086 (dd, 1 H, J = 6.2, 3.2 Hz), 4.344 (m, 1 H), 3.567 (m, 2 H), 2.686 (d, 4 H, J = 13.5 Hz), 2.030 (d, 2 H, J = 12.5 Hz), 1.78 (m, 4 H), 1.540 (s, 3 H), 1.51 (m, 4 H), 1.326 (s, 3 H), 1.35-1.1 (m, 22 H), 0.876 (t, 12 H, J = 7.0 Hz), 0.807 (t, 6 H, J = 7.0 Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₇₄H₉₀N₁₁O₁₀ (M+H), 1292.6872; found, 1292.6860.

N-(4-(4'-Methoxycarbonylphenyl)phenyl)-3,6-bis(pivaloylamino)carbazole (**C16**)

Pyridine (23 mL) was added to amine **C12** (1.1 mmol), trimethylacetyl chloride (350 μ L, 2.6 equiv), and DMAP (catalytic) in a dry round bottom flask under argon. The resulting solution was stirred at room temperature for 21 h under a drying tube, then taken up in EtOAc (100 mL). The solution was washed with water (40 mL), 1 N HCl (3 x 70 mL), water (30 mL), and brine. After drying over Na₂SO₄, solvent was removed by rotary evaporation. The residue was purified by flash chromatography on silica (10% EtOAc/CH₂Cl₂) followed by trituration in MeOH (8 mL) to give 256 mg (40% yield) of off-white powder: IR (KBr) 3270, 2954, 1725, 1646, 1529, 1491, 1466, 1279, 1208, 1180, 1113, 807, 773 cm⁻¹; ¹H NMR (500 MHz, DMSO-d₆) δ 9.296 (s, 2 H, amide), 8.485 (d, 2 H, J = 2.0 Hz), 8.095 (d, 2 H, J = 8.5 Hz), 8.036 (d, 2 H, J = 8.5 Hz), 7.953 (d, 2 H, J = 8.5 Hz), 7.755 (d, 2 H, J = 8.5 Hz), 7.596 (dd, 2 H, J = 9.0, 2.0 Hz), 7.426 (d, 2 H, J = 9.0 Hz), 3.892 (s, 3 H), 1.270 (s, 18 H); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₃₆H₃₈N₃O₄ (M+H), 576.2862; found, 576.2859.

N-(4-(4'-Carboxyphenyl)phenyl)-3,6-bis(pivaloylamino)carbazole (**C17**)

Dealkylation of methyl ester **C16** was performed in the same manner as for ester **C13** to give quantitative yield of acid **C17** which was used without further purification: ¹H NMR (300 MHz, DMSO-d₆) δ 9.299 (s, 2 H, amide), 8.488 (d, 2 H, J = 1.5 Hz), 8.074 (d, 2 H, J = 8.4 Hz), 8.026 (d, 2 H, J = 8.7 Hz), 7.919 (d, 2 H, J = 8.4 Hz), 7.748 (d, 2 H, J = 8.7 Hz), 7.596 (dd, 2 H, J = 9.0, 1.8 Hz), 7.426 (d, 2 H, J = 9.0 Hz), 1.270 (s, 18 H).

Control template (**C18**)

Conversion of acid **C17** (51 mg, 0.091 mmol) to the acid chloride, followed by condensation to with 5'-amino-5'-deoxy-2',3'-isopropylideneadenosine (1.08 equiv) was performed according to the procedure for template **C15** to yield 41 mg (53%) of off-white powder following chromatography on silica (4.5% MeOH/CHCl₃): IR (KBr) 3342, 2959, 1647, 1609, 1590, 1533, 1491, 1468, 1329, 1296, 1210, 1093, 866, 798 cm⁻¹; ¹H NMR (500 MHz, DMSO-d₆) δ 9.286 (s, 2 H), 8.737 (br t, 1 H, amide), 8.482 (d, 2 H, J = 2.0 Hz), 8.343 (s, 1 H), 8.098 (s, 1 H), 8.021 (d, 2 H, J = 9.0 Hz), 7.981 (d, 2 H, J = 8.5 Hz), 7.888 (d, 2 H, J = 8.5 Hz), 7.738 (d, 2 H, J = 8.5 Hz), 7.595 (dd, 2 H, J = 9.0, 2.0 Hz), 7.421 (d, 2 H, J = 9.0 Hz), 6.170 (d, 1 H, J = 2.5 Hz), 5.512 (dd, 1 H, J = 6.2, 2.7 Hz), 5.087 (dd, 1 H, J = 6.2, 3.2 Hz), 4.338 (q, 1 H, J = 3.5 Hz), 3.568

(m, 2 H), 1.270 (s, 18 H); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₄₈H₅₂N₉O₆ (M+H), 850.4041; found, 850.4033.

N-(4-(4'-(Methoxycarbonyl)phenyl)phenyl)-3,6-bis(((cis,cis-2,4-dioxo-1,5,7-tris((benzyloxy)methyl)-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino)carbazole (**C19**)

In a manner analogous to cleft **C13**, tris(benzyloxymethyl) Kemp's acyl chloride⁴² (3.62 mmol) and diamine **C12** (1.816 mmol) were refluxed in pyridine (10 mL) with catalytic DMAP to give 996 mg (37%) of beige solid, following flash chromatography (15% EtOAc/CH₂Cl₂): ¹H NMR (500 MHz, DMSO-d₆) δ 10.643 (s, 2 H, imide), 9.410 (s, 2 H, amide), 8.326 (d, 2 H, J = 1.0 Hz), 8.099 (d, 2 H, J = 8.5 Hz), 8.050 (d, 2 H, J = 8.5 Hz), 7.966 (d, 2 H, J = 8.5 Hz), 7.748 (d, 2 H, J = 8.5 Hz), 7.488 (dd, 2 H, J = 9.0, 1.0 Hz), 7.393 (d, 2 H, J = 8.5 Hz), 7.36-7.23 (m, 30 H), 4.518 (s, 4 H), 4.509 (s, 4 H), 4.466 (s, 4 H), 3.893 (s, 3 H), 3.784 (d, 4 H, J = 9.0 Hz), 3.464 (s, 4 H), 3.415 (d, 4 H, J = 9.5 Hz), 2.582 (dd, 4 H, J = 14.5, 0.5 Hz), 2.342 (dd, 2 H, J = 12.5, 0.5 Hz), 1.613 (d, 4 H, J = 14.0 Hz), 1.597 (d, 2 H, J = 13.0 Hz).

N-(4-(4'-Carboxyphenyl)phenyl)-3,6-bis(((cis,cis-2,4-dioxo-1,5,7-tris((benzyloxy)methyl)-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino)carbazole (**C20**).

The ester **C19** (520 mg, 0.347 mmol) and NaMe (6.2 equiv) were stirred in anhydrous DMSO (8 mL) under Ar for 1 h at 80 °C. The solution was cooled and taken up in CHCl₃ (50 mL). The organic layer was washed with 1 N HCl (25 mL), sat'd NH₄Cl (2 x 25 mL), and brine (50 mL), and dried over Na₂SO₄. The combined organic layer was concentrated and purified by chromatography (5% MeOH/CHCl₃) to give 362 mg (70%) of pale yellow acid **C20**: IR (KBr) 3369, 3179, 2858, 1700, 1466, 1099, 735, 698 cm⁻¹; ¹H NMR (500 MHz, DMSO-d₆) δ 10.640 (s, 2 H, imide), 9.409 (s, 2 H, amide), 8.327 (d, 2 H, J = 1.7 Hz), 8.068 (d, 2 H, J = 8.3 Hz), 8.024 (d, 2 H, J = 8.5 Hz), 7.886 (d, 2 H, J = 8.1 Hz), 7.729 (d, 2 H, J = 8.5 Hz), 7.486 (dd, 2 H, J = 8.9, 1.6 Hz), 7.388 (d, 2 H, J = 8.8 Hz), 7.36-7.23 (m, 30 H), 4.519 (s, 4 H), 4.509 (s, 4 H), 4.467 (s, 4 H), 3.786 (d, 4 H, J = 9.0 Hz), 3.466 (s, 4 H), 3.417 (d, 4 H, J = 8.8 Hz), 2.585 (d, 4 H, J = 13.9 Hz), 2.344 (d, 2 H, J = 13.7 Hz), 1.619 (d, 4 H, J = 12.5 Hz), 1.599 (d, 2 H, J = 12.7 Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for C₉₁H₈₆N₅O₁₄ (M+H), 1472.6171; found, 1472.6154.

Tris((benzyloxy)methyl) biphenylscorpion template (**C21**)

In a manner analogous to the formation of **C15**, acid **C20** (135 mg, 0.091 mmol) was converted to the acid chloride with oxalyl chloride (50 μL, 6.3 equiv) and catalytic DMF (1 μL) in anhydrous THF (10 mL). 5'-Amino-5'-deoxy-2',3'-(isopropylidene)adenosine¹³⁰ (1.11 equiv) was added to the reaction vessel and anhydrous THF (12 mL) was added to dissolve the reagents. Triethylamine (50 μL, 3.94 equiv) was added to the reaction solution accompanied by immediate formation of a white precipitate. After stirring at room temperature under Ar for 20 h, the reaction was filtered and solvents removed by rotary evaporation. Flash chromatography (4% MeOH/CHCl₃) gave 117 mg (73%) of beige powder: IR (KBr)

3333, 3200, 2859, 1701, 1648, 532, 1493, 1467, 1364, 1220, 1207, 1098, 737, 698 cm^{-1} ; $^1\text{H NMR}$ (300 MHz, DMSO-d_6) δ 10.63 (s, 2H, imide), 9.40 (s, 2 H, amide), 8.75 (br t, 1 H, amide), 8.369 (s, 1 H), 8.323 (d, 2 H, $J = 1.2$ Hz), 8.123 (s, 1 H), 8.031 (d, 2 H, $J = 8.7$ Hz), 7.982 (d, 2 H, $J = 8.1$ Hz), 7.896 (d, 2 H, $J = 8.1$ Hz), 7.725 (d, 2 H, $J = 8.4$ Hz), 7.486 (d, 2 H, $J = 8.4$), 7.4-7.2 (m, 32 H), 6.177 (d, 1 H, $J = 2.7$ Hz), 5.508 (dd, 1 H, $J = 5.8, 2.8$ Hz), 5.086 (dd, 1 H, $J = 6.9, 3.0$ Hz), 4.514 (s, 8 H), 4.467 (s, 4 H), 4.342 (m, 1 H), 3.786 (d, 4 H, $J = 8.8$ Hz), 3.55 (m, 2 H), 3.466 (s, 4 H), 3.418 (d, 4 H, $J = 9$ Hz), 2.584 (d, 4 H, $J = 14.1$ Hz), 2.345 (d, 2 H, $J = 13.5$ Hz), 1.598 (m, 6 H), 1.541 (s, 3 H), 1.334 (s, 3 H); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for $\text{C}_{104}\text{H}_{102}\text{N}_{12}\text{O}_{16}$ (M+H), 1760.7506; found, 1760.7491.

Tris(hydroxymethyl) biphenylscorpion template (C22)

The protected template **C21** (24 mg, 0.0136 mmol) was dissolved in HCOOH (4.8 mL) and stirred on ice for 0.5 h. HBr(g) was bubbled through the solution for 8 min followed by Ar for 2 h. Solvent was removed by rotary evaporation under reduced pressure. The residue was suspended in MeOH (< 1 mL) and precipitated with Et_2O (2 mL). After filtering, the solid was triturated in THF (2 mL) for 1 h, filtered and dried under vacuum to give 11.6 mg (73%) of beige powder: IR (KBr) 3331, 2938, 2874, 1692, 1641, 1609, 1529, 1494, 1466, 1202, 1056, 811, 772 cm^{-1} ; $^1\text{H NMR}$ (500 MHz, DMSO-d_6) δ 10.377 (s, 2 H, imide), 9.230 (s, 2 H, amide), 8.771 (s, 1 H), 8.7331 (br t, 1 H, amide), 8.5 (s, 2 H, amine), 8.451 (s, 1 H), 8.361 (s, 2 H), 8.029 (d, 2 H, $J = 8.5$ Hz), 8.000 (d, 2 H, $J = 8.0$ Hz), 7.905 (d, 2 H, $J = 8.5$ Hz), 7.723 (d, 2 H, $J = 8.5$ Hz), 7.494 (d, 2 H, $J = 9.0$ Hz), 7.367 (d, 2 H, $J = 8.5$ Hz), 5.964 (d, 1 H, $J = 6.0$ Hz), 4.700 (t, 1 H, $J = 5.5$ Hz), 4.54 (br, 6 H, hydroxyl), 4.217 (m, 1 H), 4.149 (q, 1 H, $J = 3.5$ Hz), 3.779 (d, 4 H, $J = 10.5$ Hz), 3.66 (m, 2 H), 3.404 (s, 4 H), 3.336 (d, 4 H, $J = 10.5$ Hz), 2.419 (d, 4 H, $J = 14.0$ Hz), 2.203 (d, 4 H, $J = 13.0$ Hz), 1.488 (d, 4 H, $J = 14.0$ Hz), 1.341 (d, 2 H, $J = 13.0$ Hz); HRMS (FAB in 3-nitrobenzyl alcohol) calcd for $\text{C}_{59}\text{H}_{62}\text{N}_{11}\text{O}_{16}$ (M+H), 1180.4376; found, 1180.4379.

N-(4-(4'-Carboxyphenyl)phenyl)-3,6-bis(((cis,cis-2,4-dioxo-1,5,7-tris(hydroxymethyl)-3-azabicyclo[3.3.1]non-7-yl)carbonyl)amino)carbazole (C23)

HBr(g) was bubbled through a suspension of acid **C20** (362 mg, 0.244 mmol) in HCOOH (25 mL) at 0°C for 10 min. Solvent was removed by rotary evaporation under reduced pressure and the residue pumped dry (KOH trap). The crude material was triturated in CHCl_3 and filtered. The resulting orange solid (225 mg) was suspended in MeOH (12 mL) and pure white product (82.9 mg, 39%) was precipitated with ether. The filtrate was concentrated and triturated in THF (1.5 mL) to give slightly less pure compound as a beige powder (84.3 mg, 40%) after filtration and drying: IR (KBr) 3363, 2940, 2878, 1695, 1527, 1467, 1202, 1055 cm^{-1} ; $^1\text{H NMR}$ (500 MHz, DMSO-d_6) δ 10.408 (s, 2 H, imide), 9.242 (s, 2 H, amide), 8.372 (d, 2 H, $J = 2.0$ Hz), 8.073 (d, 2 H, $J = 8.3$ Hz), 8.035 (d, 2 H, $J = 8.5$ Hz), 7.928 (d, 2 H, $J = 8.3$ Hz), 7.733 (d, 2 H, $J = 8.5$ Hz), 7.490 (dd, 2 H, $J = 8.9, 1.8$ Hz), 7.381 (d, 2 H, $J = 8.8$ Hz), 3.781 (d, 4 H, $J = 10.3$ Hz), 3.38 (d, 8 H, $J = 12$ Hz), 2.417 (d, 4 H, $J = 14.9$ Hz), 2.205 (d, 2 H, $J = 12.5$ Hz), 1.484 (d, 4 H, $J = 14.9$ Hz),

1.336 (d, 2 H, $J = 13.2$ Hz); HRMS (FAB in glycerol) calcd for $C_{49}H_{50}N_5O_{14}$ (M+H), 932.3354; found, 932.3359.

N-(4-(4'-Carboxyphenyl)phenyl)-3,6-bis[*[(cis,cis-2,4-dioxo-1,5,7-tris(acetoxymethyl)-3-azabicyclo[3.3.1]non-7-yl)carbonyl]amino*]carbazole (**C24**).

Benzyl-protected acid **C20** (35.2 mg, 0.0237 mmol) was dissolved in HOAc (5 mL) and Ac_2O (0.4 mL, 178 equiv). HBr(g) was bubbled through the solution for 10 min followed by Ar for 30 min as purge. Solvent was removed by rotary evaporation under reduced pressure and under vacuum (KOH trap) and the residue purified by flash chromatography (5-100% MeOH/ $CHCl_3$) to give an off-white solid identified as partially-pure product: 1H NMR (500 MHz, $DMSO-d_6$) δ 10.944 (s, 2 H, imide), 9.569 (s, 2 H, amide), 8.238 (d, 2 H, $J = 1.5$ Hz), 8.074 (d, 2 H, $J = 8.0$ Hz), 8.038 (d, 2 H, $J = 8.5$ Hz), 7.925 (d, 2 H, $J = 8.0$ Hz), 7.739 (d, 2 H, $J = 8.5$ Hz), 7.480 (dd, 2 H, $J = 8.5, 2.0$ Hz), 7.406 (d, 2 H, $J = 9.0$ Hz), 4.353 (d, 2 H, $J = 11.0$ Hz), 4.131 (s, 2 H), 4.056 (d, 2 H, $J = 11.0$ Hz), 2.684 (d, 4 H, $J = 14.0$ Hz), 2.239 (d, 2 H, $J = 13.0$ Hz), 2.033 (s, 12 H), 2.007 (s, 6 H), 1.604 (d, 4 H, $J = 14.0$ Hz), 1.565 (d, 2 H, $J = 13.0$ Hz); HRMS (FAB in glycerol) calcd for $C_{61}H_{61}N_5O_{20}$, 1183.3910; found, 1183.3911.

5.2 Spectroscopy

Dilution and variable-temperature 1H NMR studies were performed at 500 MHz using Varian VXR-500 NMR spectrometers equipped with Oxford VTC-4 temperature controllers and either 5 mm direct or indirect detection $^1H/^{19}F$ probes. Temperatures are not calibrated. $DMSO-d_6$ (99.9%-D), THF- d_8 (99.5%-D), and "100%" $CDCl_3$ (99.96%-D) were supplied by Cambridge Isotope Labs.

5.3 Titrations

1H titrations were performed at 298 K at 300 MHz on a Varian UN-300 instrument with a 5 mm $^1H/^{19}F$ /broad band probe and an Oxford VTC-4 temperature controller (THF) or at 500 MHz on a Varian VXR-500 instrument with a 5 mm $^1H/^{19}F$ indirect detection probe and an Oxford VTC-4 temperature controller ($CDCl_3$).

In THF- d_8 (99.5%-D, CIL), the imide proton was followed through a 3.22 ppm downfield change in chemical shift as 2',3'-isopropylideneadenosine (10 mM) was added to a 2 mM solution of biphenylscorpion **C13** (700 μ L). In $CDCl_3$ (99.8%-D, CIL), the upfield shift of adenine aromatic protons was monitored as a solution of **C13** (2 mM) was added to 2',3'-isopropylidene (1 mM).

Association constants were calculated by non-linear least-squares regression fit of the data to the 1:1 binding isotherm using Systat 5.2. The final chemical shift and association constant were treated as variables; the experimental value for the initial chemical shift was used.

5.4 Molecular Modeling

Predicted structures are generated by molecular mechanics minimization on a Silicon Graphics 4D30G+ Personal Iris running MacroModel 3.5X⁷⁰ using a modified AMBER⁷² force field (AMBER*) and the truncated Newton conjugate gradient algorithm⁹⁴ (TNCG) to the default gradient of less than 0.05 kJ/Å. Chloroform solution was simulated using the Generalized Born/Surface Area method (GB/SA).⁷³

5.5 Vapor Phase Osmometry

All molecular weight determinations were performed by Ivan Huc at 30 °C with a Wescan Model 233 vapor pressure osmometer in the laboratory of George Whitesides at Harvard University. Standards were measured at 1, 3, 6, and 10 mM and molecular weights determined at 3 mM in HPLC grade glass-distilled chloroform from Aldrich. There is approximately a $\pm 20\%$ margin of error for the measurements.

5.6 Kinetics

All kinetic measurements were performed by Edward Wintner. Reaction rates for reaction between active ester derivatives of **C14** and 5'-amino-5'-deoxy-2',3'-isopropylideneadenosine¹³⁰ were measured by HPLC monitoring of product formation at 266 nm with a Waters 490E multiwavelength detector run by a 600E System controller equipped with 717 Autosampler and controlled by Baseline 810 HPLC software or by monitoring ester consumption (2,4-dinitrophenyl ester) at 450 nm with a Perkin-Elmer Lambda2 UV/Visible spectrometer controlled by PECCS UV/Vis software at room temperature (23 °C).

D. BIBLIOGRAPHY

- (1) Rebek, J., Jr. *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 245-255.
- (2) Gowsami, S.; Hamilton, A. D.; Van Engen, D. *J. Am. Chem. Soc.* **1989**, *111*, 3425-3426.
- (3) Wilcox, C. S. In *Supramolecular Organic Chemistry and Photochemistry*; H.-J. Schneider and H. Durr, Ed.; VCH: Weinheim, 1990.
- (4) Zimmerman, S. C.; Wu, W.; Zeng, Z. *J. Am. Chem. Soc.* **1991**, *113*, 196-201.
- (5) Park, T. K.; Schroeder, J.; Rebek, J., Jr. *Tetrahedron* **1991**, *47*, 2507-2518.
- (6) Park, T. K.; J., S.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 5125-5127.
- (7) Hosseini, M. W.; Blacker, A. J.; Lehn, J.-M. *J. Am. Chem. Soc.* **1990**, *112*, 3896-3904.
- (8) Furata, H.; Magda, D.; Sessler, J. L. *J. Am. Chem. Soc.* **1991**, *113*, 978-985.
- (9) Goodman, M. S.; Rose, S. D. *J. Am. Chem. Soc.* **1991**, *113*, 9380-9382.
- (10) Inouye, M.; Kim, K.; Kitao, T. *J. Am. Chem. Soc.* **1992**, *114*, 778-780.
- (11) Ts'o, P. O. P. In *Basic Principles in Nucleic Acid Chemistry*; P. O. P. Ts'o, Ed.; Academic Press: New York, 1974; Vol. I; pp 454-584.
- (12) Kyogoku, Y.; Lord, R. C.; Rich, A. *Proc. Natl. Acad. Sci.* **1967**, *57*, 250-257.
- (13) Kyogoku, Y.; Lord, R. C.; Rich, A. *Biochim. Biophys. Acta* **1969**, *179*, 10-17.
- (14) Williams, K.; Askew, B.; Ballester, P.; Buhr, C.; Jeong, K. S.; Jones, S.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1989**, *111*, 1090-1094.
- (15) Jeong, K. S.; Tjivikua, T.; Muehldorf, A.; Deslongchamps, G.; Famulok, M.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 201-209.
- (16) Benzing, T.; Tjivikua, T.; Wolfe, J.; Rebek, J., Jr. *Science* **1988**, *242*, 266-268.
- (17) Fersht, A. R. *Trends Biochem. Sci.* **1987**, *12*, 301-304.
- (18) Galán, A.; de Mendoza, J.; Toiron, C.; Bruix, M.; Deslongchamps, G.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 9424-9425.
- (19) Deslongchamps, G.; Galán, A.; de Mendoza, J.; Rebek, J., Jr. *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 61-63.
- (20) Conn, M. M.; Deslongchamps, G.; de Mendoza, J.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1993**, *115*, 3548-3557.
- (21) Olsher, U.; Frolow, F.; Dalley, N. K.; Weiming, J.; Yu, Z.-Y.; Knobloch, J. M.; Bartsch, R. A. *J. Am. Chem. Soc.* **1991**, *113*, 6570-6574.
- (22) Jeong, K. S.; Muehldorf, A. V.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1990**, *112*, 6144-6145.
- (23) for a review of this reaction, see: Mitsunobu, O. *Synthesis* **1981**, 1-28.
- (24) Fersht, A. R.; Shi, J.-P.; Knill-Jones, J.; Lowe, D. M.; Wilkinson, A. J.; Blow, D. M.; Brick, P.; Carter, P.; Waye, M. M. Y.; Winter, G. *Nature* **1985**, *314*, 235-238.
- (25) Williams, D. H.; Cox, J. P. L.; Doig, A. J.; Gardner, M.; Gerhard, U.; Kaye, P. T.; Lal, A. R.; Nicholls, I. A.;

- Salter, C. J.; Mitchell, R. C. *J. Am. Chem. Soc.* **1991**, *113*, 7020-7030.
- (26) Connors, K. A. *Binding Constants*; 1st ed.; John Wiley & Sons: New York, 1987, pp 411.
- (27) Sandström, J. *Dynamic NMR Spectroscopy*; Academic Press: London, 1982.
- (28) Braunschweiler, L.; Ernst, R. R. *J. Magn. Reson.* **1983**, *53*, 521-528.
- (29) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546-4553.
- (30) Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; VCH: New York, 1989.
- (31) Bothner-By, A. A.; Stephens, R. L.; Lee, J. *J. Am. Chem. Soc.* **1984**, *106*, 811-813.
- (32) Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *63*, 355-360.
- (33) Neuhaus, D.; Keeler, J. *J. Magn. Reson.* **1986**, *68*, 568-574.
- (34) Kessler, H.; Griesinger, C.; Kerssebaum, R.; Wagner, K.; Ernst, R. R. *J. Am. Chem. Soc.* **1987**, *109*, 607-609.
- (35) Systat 5.2, Systat Inc., Evanston, Ill, 1992.
- (36) Ghislain Deslongchamps, unpublished results.
- (37) Schneider, H.-J. *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 1417-1436.
- (38) Gerhard, U.; Searle, M. S.; Williams, D. H. *Biorg. Med. Chem. Lett.* **1993**, *3*, 803-808.
- (39) Adrian, J. C., Jr.; Wilcox, C. S. *J. Am. Chem. Soc.* **1989**, *111*, 8055-8057.
- (40) Hore, P. J. *J. Magn. Reson.* **1983**, *54*, 539-542.
- (41) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984, pp 556.
- (42) Rotello, V. M.; Viani, E. A.; Deslongchamps, G.; Murray, B. A.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1993**, *115*, 797-798.
- (43) Kato, Y.; Conn, M. M.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1994**, *116*, in press.
- (44) Linse, P. *J. Am. Chem. Soc.* **1992**, *114*, 4366-4373.
- (45) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, 1991.
- (46) Williams, D. H.; Searle, M. S.; Mackay, J. P.; Gerhard, U.; Maplestone, R. A. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1172-1178.
- (47) Searle, M. S.; Williams, D. H. *An. Quim.* **1993**, *89*, 17-26.
- (48) Page, M. I.; Jencks, W. P. *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 1678-1683.
- (49) Jencks, W. P. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 4046-4050.
- (50) Pilichowski, J.; Puszynski, A. *Monatsh. Chem.* **1974**, *105*, 772-774.
- (51) Vincent Rotello, Yoko Kato, Doris Lee, Belinda Tsao, unpublished data.
- (52) Kyu-Sung Jeong, unpublished data.
- (53) Brian A. Murray, personal communication.
- (54) Kneeland, D. M.; Ariga, K.; Lynch, V. M.; Huang, C.-Y.; Anslyn, E. V. *J. Am. Chem. Soc.* **1993**, *115*, 10042-10055.

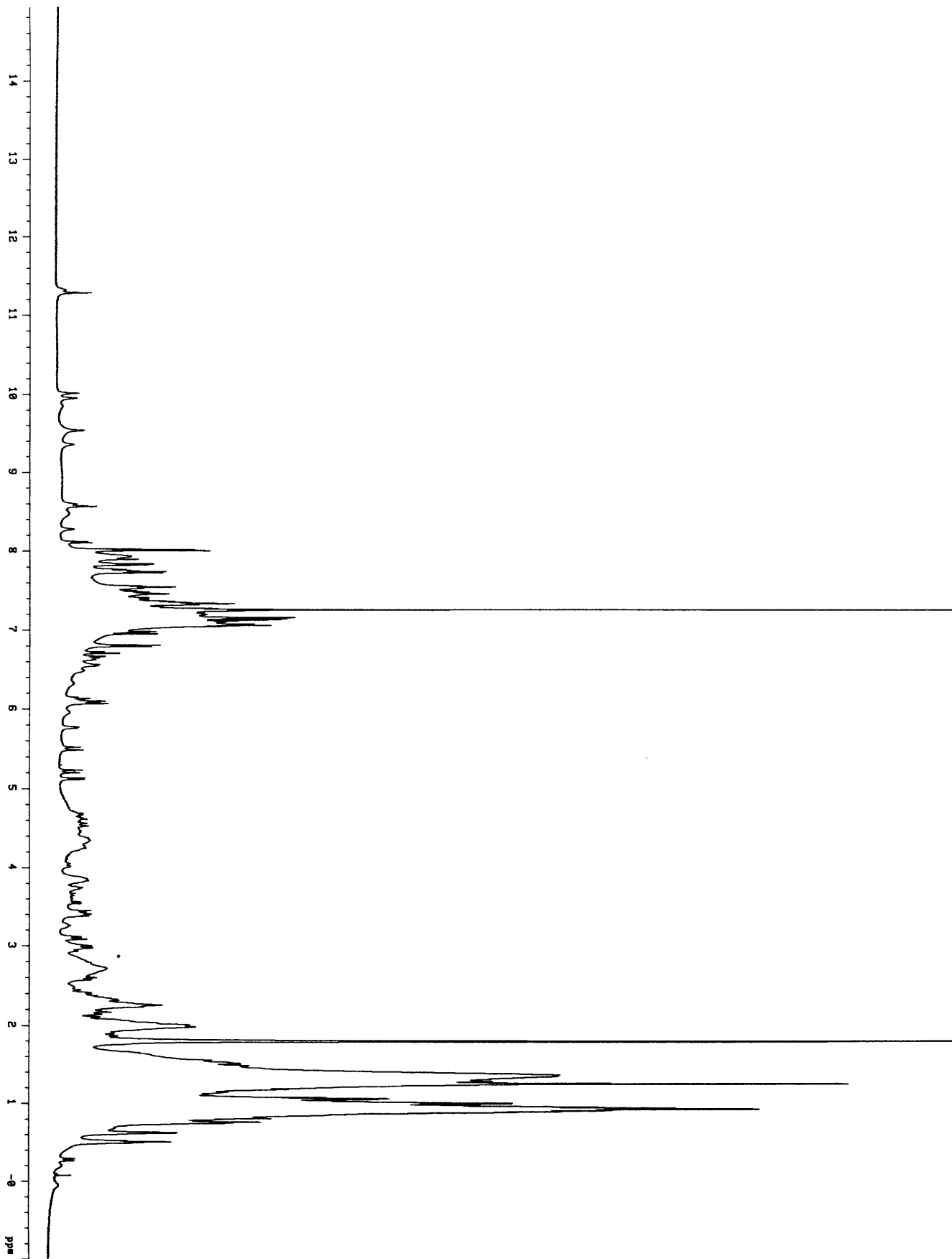
- (55) Belinda Tsao, personal communication.
- (56) Askew, B.; Ballester, P.; Buhr, C.; Jeong, K. S.; Jones, S.; Parris, K.; Williams, K.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1989**, *111*, 1082-1090.
- (57) Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990**, *112*, 5525-5534.
- (58) Muehldorf, A. V.; Van Engen, D.; Warner, J. c.; Hamilton, A. D. *J. Am. Chem. Soc.* **1988**, *110*, 6561-6562.
- (59) Still, W. C.; Kahn, M.; Mitra, A. *J. Org. Chem.* **1978**, *43*, 2923-2925.
- (60) Grotta, H. M.; Riggle, C. J.; Bearnse, A. E. *J. Org. Chem.* **1964**, *29*, 2474-2476.
- (61) Fieser, L. F.; Fieser, M. *Reagents for Organic Synthesis*; Wiley: New York, 1967.
- (62) Drobny, G.; Pines, A.; Sinton, S.; Weitekamp, D.; Wemmer, D. *Faraday Div. Chem. Soc. Symp.* **1979**, *13*, 49.
- (63) States, D. J.; Haberkorn, R. A.; Ruben, D. J. *J. Magn. Reson.* **1982**, *48*, 286-292.
- (64) Piantini, U.; Sørensen, O. W.; Ernst, R. R. *J. Am. Chem. Soc.* **1982**, *104*, 6800-6801.
- (65) Rance, M.; Sørensen, O. W.; Bodenhausen, G.; Wagner, G.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Comm.* **1983**, *117*, 479-485.
- (66) Kumar, A.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Comm.* **1980**, *95*, 1-6.
- (67) Kumar, A.; Wagner, G.; Ernst, R. R.; Wüthrich, K. *J. Am. Chem. Soc.* **1981**, *103*, 3654-3658.
- (68) Hore, P. J. *J. Magn. Reson.* **1983**, *54*, 283-300.
- (69) Bax, A.; Davis, D. G. *J. Magn. Res.* **1985**, *63*, 207-213.
- (70) MacroModel 3.5X: Mohamadi, F.; Richards, N. G.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440-467.
- (71) van Gunsteren, M. F.; Berendsen, H. J. C. *Mol. Simul.* **1988**, *1*, 173.
- (72) Weiner, S. J.; Kollman, P. A.; Case, D.; Singh, U. C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765-784.
- (73) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127-6129.
- (74) Hosseini, M. W.; Lehn, J.-M. *J. Chem. Soc., Chem. Commun.* **1991**, 451-453.
- (75) Echavarren, A.; Galán, A.; de Mendoza, J.; Lehn, J.-M. *Helv. Chim. Acta* **1988**, *71*, 685-693.
- (76) Schmidtchen, F. P. *Tet. Lett.* **1989**, *30*, 4493-4496.
- (77) Furuta, H.; Cyr, M. J.; Sessler, J. L. *J. Am. Chem. Soc.* **1991**, *113*, 6677-6678.
- (78) Dixon, R. P.; Geib, S. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **1992**, *114*, 365-366.
- (79) Schneider, H.-J.; Blatter, T.; Palm, B.; Pfingstag, U.; Rüdiger, V.; Theis, I. *J. Am. Chem. Soc.* **1992**, *114*, 7704-7708.
- (80) Tabushi, I.; Kobuke, Y.; Imuta, J. *J. Am. Chem. Soc.* **1980**, *102*, 1744-1745.
- (81) Tabushi, I.; Kobuke, Y.; Imuta, J. *J. Am. Chem. Soc.* **1981**, *103*, 6152-6157.
- (82) Král, V.; Sessler, J. L.; Furuta, H. *J. Am. Chem. Soc.* **1992**, *114*, 8704-8705.
- (83) Tecilla, P.; Chang, S.-K.; Hamilton, A. D. *J. Am. Chem. Soc.* **1990**, *112*, 9586-9590.

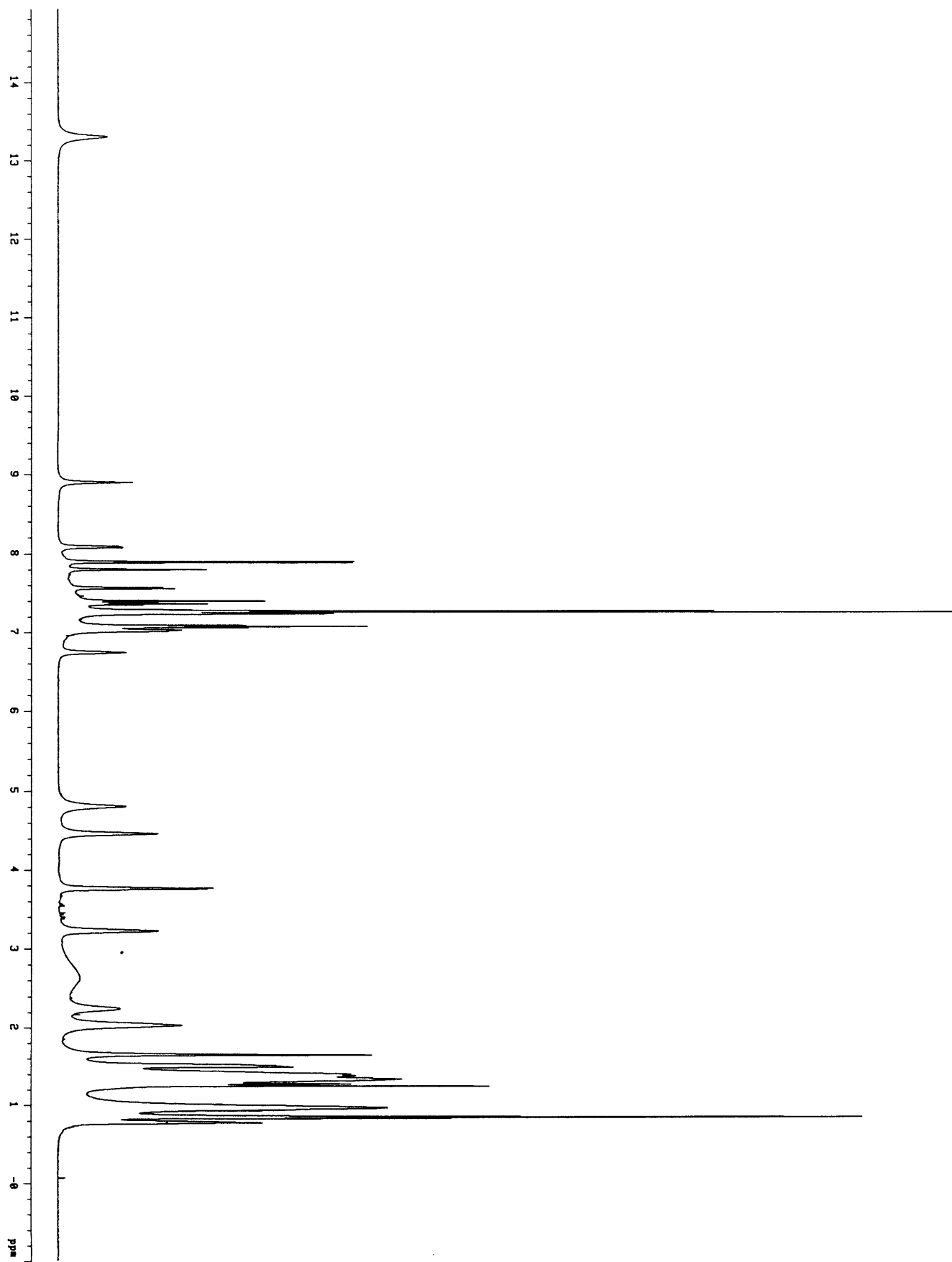
- (84) Jubian, V.; Dixon, R. P.; Hamilton, A. D. *J. Am. Chem. Soc.* **1992**, *114*, 1120-1121.
- (85) Smith, J.; Ariga, K.; Anslyn, E. V. *J. Am. Chem. Soc.* **1993**, *115*, 362-364.
- (86) for a review, see: Lindley, J. *Tetrahedron* **1984**, *40*, 1433-1456.
- (87) for a review, see: McMurry, J. *Org. React.* **1976**, *24*, 187-224.
- (88) Kurzmeier, H.; Schmidtchen, F. P. *J. Org. Chem.* **1990**, *55*, 3749-3755.
- (89) Springs, B.; Haake, P. *Bioorg. Chem.* **1977**, *6*, 181-190.
- (90) Furuta, H.; Furuta, K.; Sessler, J. L. *J. Am. Chem. Soc.* **1991**, *113*, 4706-4707.
- (91) Andreu, C.; Galán, A.; Kobiro, K.; de Mendoza, J.; Rebek, J., Jr.; Salmeron, A.; Usman, N. *J. Am. Chem. Soc.* **1994**, *116*, submitted.
- (92) Inoue, Y. In *Liquid Membranes: Chemical Applications*; T. Araki and H. Tsukube, Ed.; CRC Press: Boca Raton, 1990; pp 77-102.
- (93) Grahl, A. *Chem. Ber.* **1895**, *28*, 84-90.
- (94) Ponder, J. W.; Richards, F. M. *J. Comp. Chem.* **1987**, *8*, 1016-1024.
- (95) Levy, S. *Artificial Life: The Quest for a New Creation*; 1st ed.; Pantheon: New York, 1992, pp 390.
- (96) Miller, S. L. *Cold Spring Harbor Symp. Quant. Biol.* **1987**, *LII*, 17-27.
- (97) Drenkard, S.; Ferris, J.; Eschenmoser, A. *Helv. Chim. Acta* **1990**, *73*, 1373-1390.
- (98) Wagner, E.; Xiange, Y.-B.; Baumann, K.; Gück, J.; Eschenmoser, A. *Helv. Chim. Acta* **1990**, *73*, 1391-1409.
- (99) Müller, D.; Pitsch, S.; Kittaka, A.; Wagner, E.; Wintner, C. E.; Eschenmoser, A. *Helv. Chim. Acta* **1990**, *73*, 1410-1468.
- (100) Basile, B.; Lazcano, A.; Oró, J. *Adv. Space Res.* **1984**, *4*, 125-131.
- (101) Mar, A.; Oró, J. *J. Mol. Evol.* **1991**, *32*, 201-210.
- (102) Miller, S. L.; Schlesinger, G. *J. Mol. Evol.* **1993**, *36*, 302-307.
- (103) Miller, S. L.; Schlesinger, G. *J. Mol. Evol.* **1993**, *36*, 308-314.
- (104) Bachmann, P. A.; Luisi, P. L.; Lang, J. *Nature* **1992**, *357*, 57-59.
- (105) von Kiedrowski, G. *Angew. Chem. Int. Ed. Engl.* **1986**, *25*, 932-935.
- (106) von Kiedrowski, G.; Wlotzka, B.; Helbing, J.; Matzen, M.; Jordan, S. *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 423-426.
- (107) Achilles, T.; von Kiedrowski, G. *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 1198-1201.
- (108) Tjivikua, T.; Ballester, P.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1990**, *112*, 1249-1250.
- (109) Edward A. Wintner, unpublished results.
- (110) Titskii, G. D.; Litvinenko, L. M. *Zh. Obs. Khim.* **1970**, *70*, 2680-2688.
- (111) Menger, F. M.; Smith, J. H. *J. Am. Chem. Soc.* **1972**, *94*, 3824-3829.
- (112) Rony, P. R. *J. Am. Chem. Soc.* **1969**, *91*, 6090-6096.
- (113) Swain, C. G.; Brown, J. F., Jr. *J. Am. Chem. Soc.* **1952**, *74*, 2538-2543.

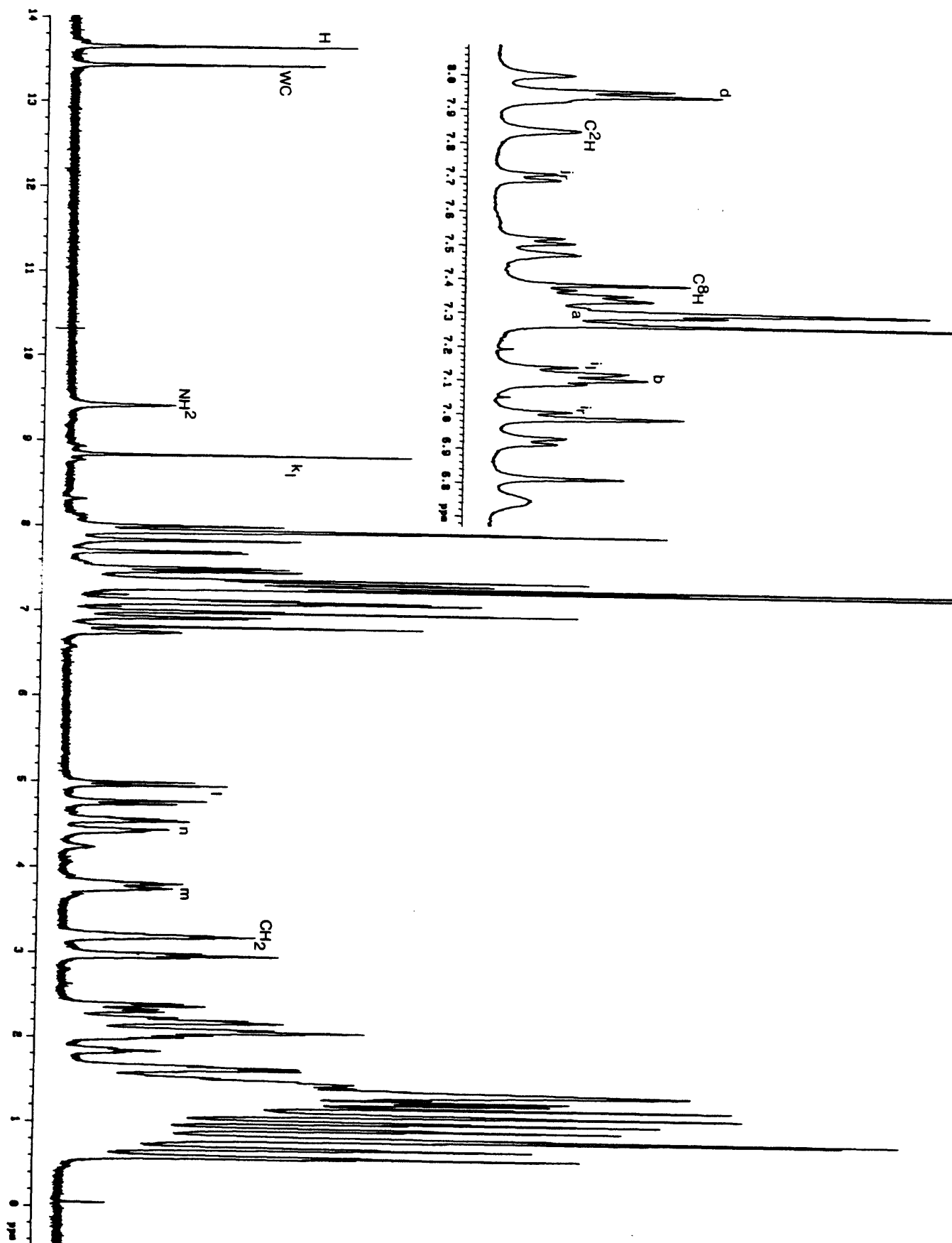
- (114) Litvinenko, L. M.; Savelova, V. A.; Skripka, A. V. *Zh. Obs. Khim.* **1970**, *40*, 886-894.
- (115) for a review of this proposal, see: Warshel, A. *Acc. Chem. Res.* **1981**, *14*, 284-290.
- (116) Nowick, J. S.; Feng, Q.; Tjivikua, T.; Ballester, P.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1993**, *113*, 8831-8839.
- (117) Rotello, V.; Hong, J. I.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 9422-9423.
- (118) Hong, J.-I.; Feng, Q.; Rotello, V.; Rebek, J., Jr. *Science* **1992**, *255*, 848-850.
- (119) Searle, M. S.; Williams, D. H. *J. Am. Chem. Soc.* **1992**, *114*, 10690-10697.
- (120) Seto, C. T.; Whitesides, G. M. *J. Am. Chem. Soc.* **1993**, *115*, 905-916.
- (121) Ivan Huc, personal communication.
- (122) ElAmin, B.; Anantharamaiah, G. M.; Royer, G. P.; Means, G. E. *J. Org. Chem.* **1979**, *44*, 3442-3444.
- (123) Yu, H.-A.; Pettitt, B. M.; Karplus, M. *J. Am. Chem. Soc.* **1991**, *113*, 2425-2434.
- (124) Stewart, D. E.; Sarkar, A.; Wampler, J. E. *J. Mol. Biol.* **1990**, *214*, 253-260.
- (125) Tweedy, N. B.; Nair, S. K.; Paterno, S. A.; Fierke, C. A.; Christianson, D. W. *Biochem.* **1993**, *32*, 10944-10949.
- (126) Jorgensen, W. L.; Severance, D. L. *J. Am. Chem. Soc.* **1991**, *113*, 209-216.
- (127) Polak, E.; Ribiere, G. *Revue Francaise Inf. Rech. Oper.* **1969**, *16*, 35.
- (128) Eigen, M.; Schuster, P. *The Hypercycle*; Springer-Verlag: New York, 1979.
- (129) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley & Sons: New York, 1986, pp 292.
- (130) Kolb, M.; Danzin, C.; Barth, J.; Claverie, N. *J. Med. Chem.* **1982**, *25*, 550-556.
- (131) Leonard, N. J.; Lambert, R. F. *J. Org. Chem.* **1969**, *34*, 3240-3248.
- (132) Merkushev, E. B.; Yudina, N. D. *Zh. Org. Khim.* **1982**, *17*, 2598-2601.

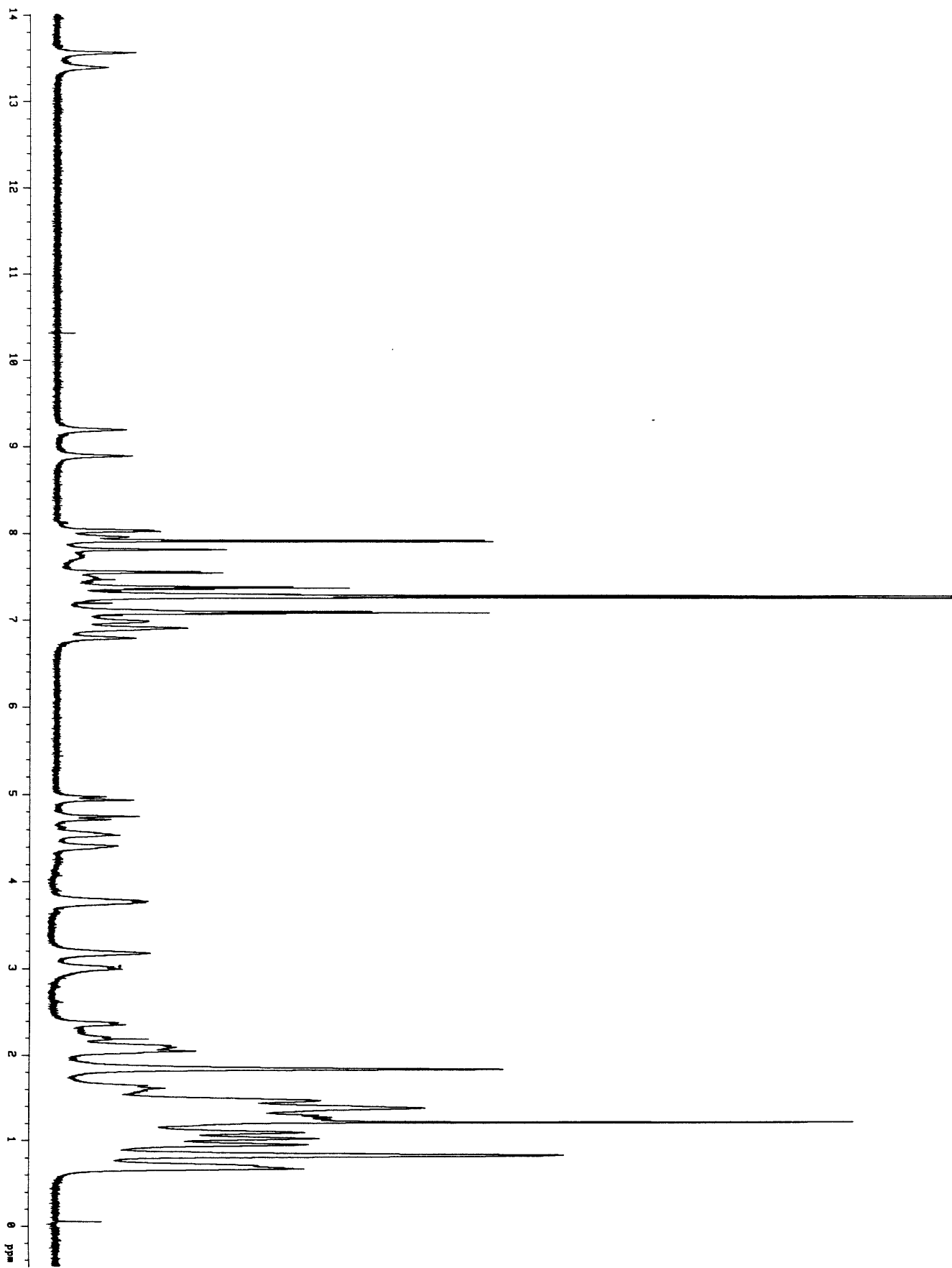
E. SPECTRA

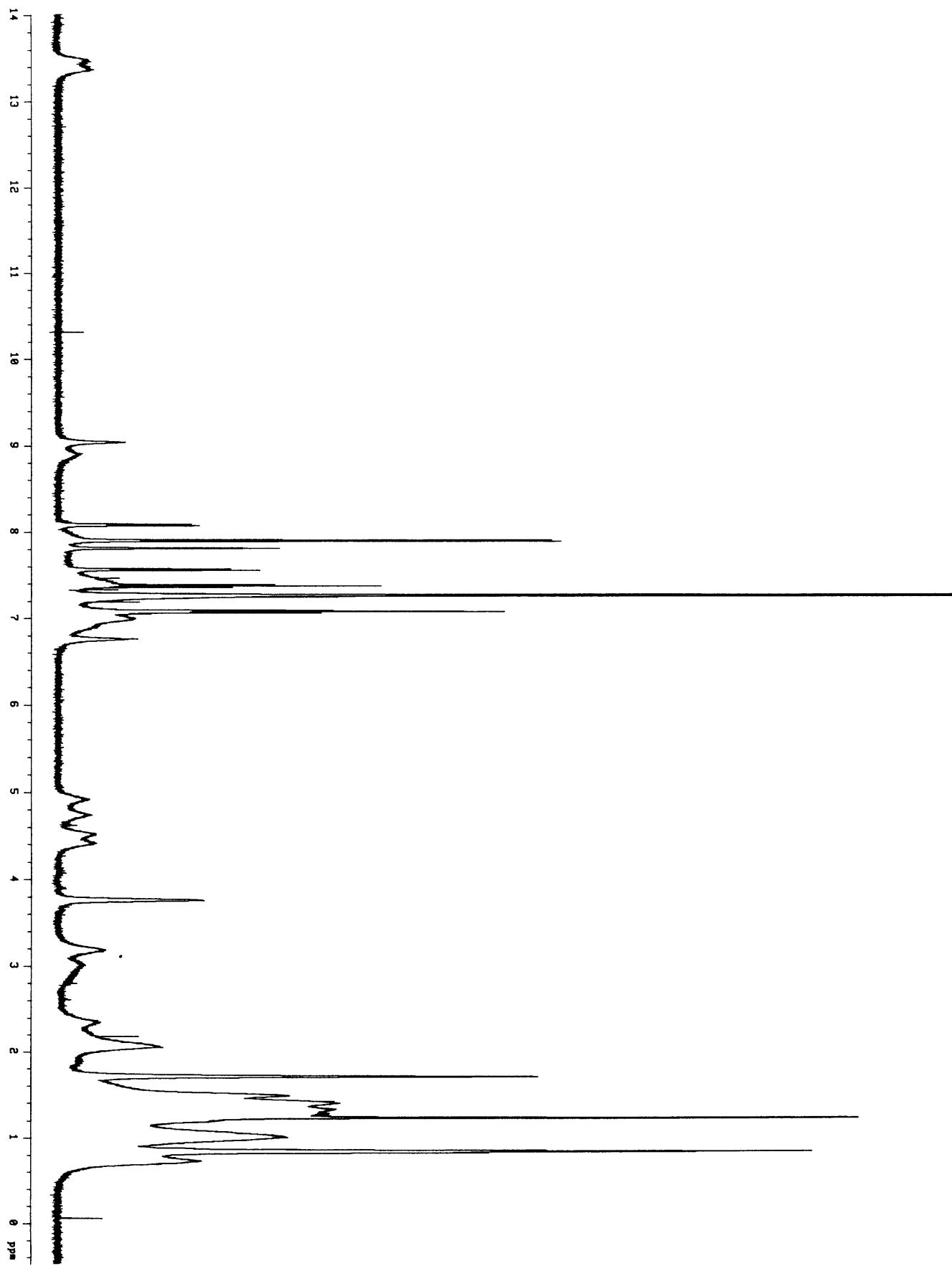
- Figure E.1 500 MHz ^1H NMR spectrum at 25 °C of **A24** (9.98 mM) in CDCl_3 (a) before and (b) after addition of 1.036 equiv of 9-ethyladenine.
- Figure E.2 500 MHz ^1H NMR spectrum of **A24** (5.47 mM) and 9-ethyladenine (5.57 mM) in CDCl_3 at (a) -50 °C, (b) -20 °C, (c) 0 °C, (d) 35 °C, and (e) 45 °C.
- Figure E.3 300 MHz ^1H NMR TOCSY spectrum at 25 °C of the complex of **A27** (5.00 mM) and 9-ethyladenine (5.01 mM) in CD_2Cl_2 .
- Figure E.4 300 MHz ^1H NMR NOESY spectrum at 25 °C of the complex of **A27** (5.00 mM) and 9-ethyladenine (5.01 mM) in CD_2Cl_2 showing (a) the full spectrum, (b) imide-aromatic correlations, and (c) aromatic-aliphatic correlations.
- Figure E.5 500 MHz ^1H NMR TOCSY spectrum at -50 °C of the complex of **A24** (10.2 mM) and 9-ethyladenine (10.4 mM) in CDCl_3 .
- Figure E.6 500 MHz ^1H NMR NOESY spectrum at -50 °C of the complex of **A24** (10.2 mM) and 9-ethyladenine (10.4 mM) in CDCl_3 showing (a) the full spectrum, (b) imide-aromatic correlations, (c) aromatic correlations, and (d) aromatic-aliphatic correlations.
- Figure E.7 500 MHz ^1H NMR spectra of **C1** (11.2 mM) in CDCl_3 at (a) 0 °C, (b) 10 °C, (c) 20 °C, (d) 25 °C, (e) 30 °C, and (f) 40 °C.
- Figure E.8 The imide region of the 500 MHz ^1H NMR spectrum at 25 °C of **C1** in CDCl_3 at varying concentrations.
- Figure E.9 500 MHz ^1H NMR spectra of **C2** in CDCl_3 at 25 °C showing (a) the full spectrum at 11.25 mM and (b) the imide region with varying concentration.
- Figure E.10 500 MHz ^1H NMR spectra of **C5** (6.7 mM) in CDCl_3 at 20 °C (a) and showing the imide region at varying temperature (b and c).
- Figure E.11 500 MHz ^1H NMR spectra of **C15** (5 mM) in CDCl_3 at 25 °C (a) and from -50 °C to 50 °C (b, imide region only).
- Figure E.12 UV/visible spectra for **C15**, **C18** and **C15** with **C18** (10 μmol of each) in 4% MeOH/ CHCl_3 . The mathematical sum of the individual spectra is also plotted.

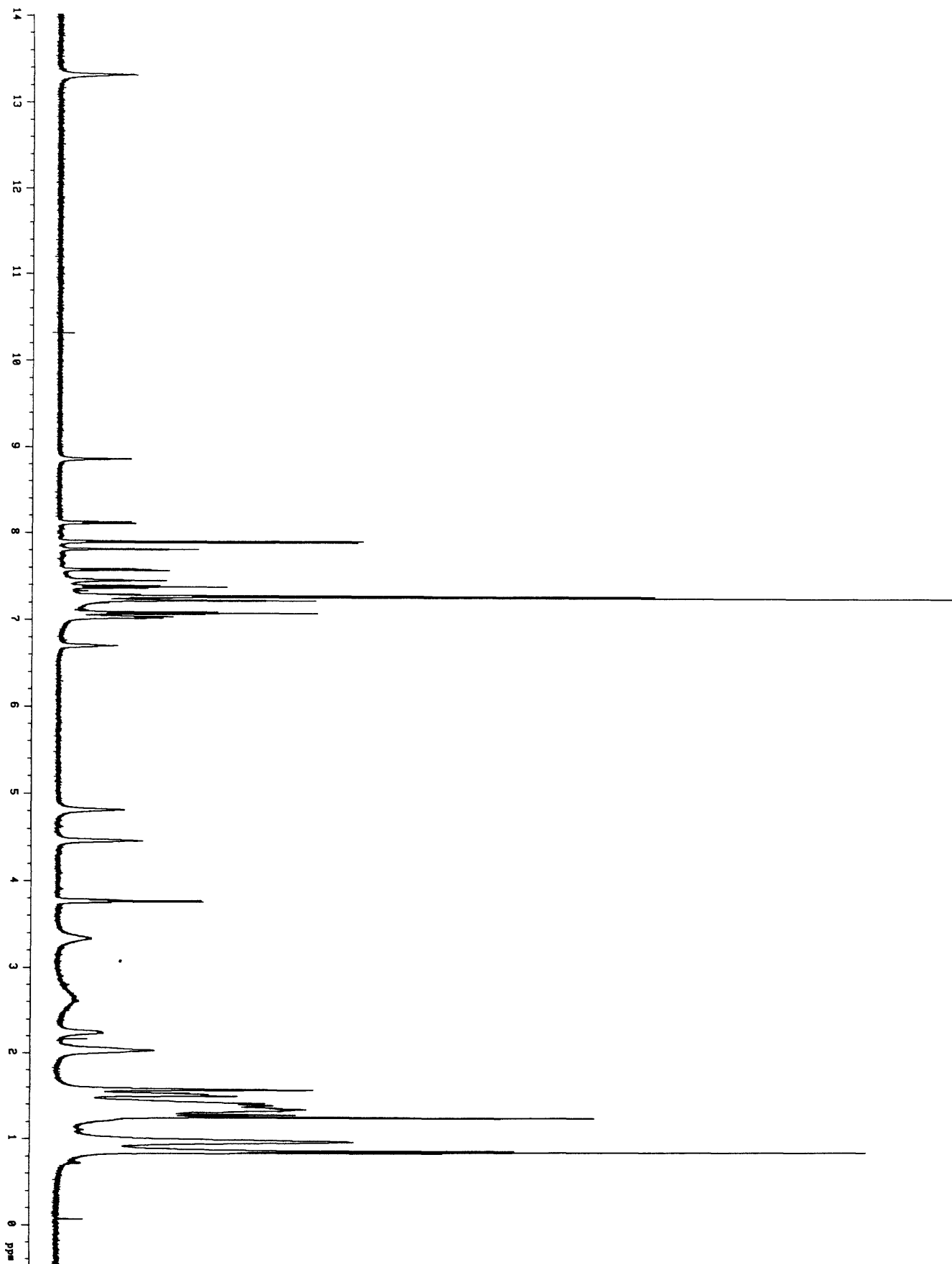


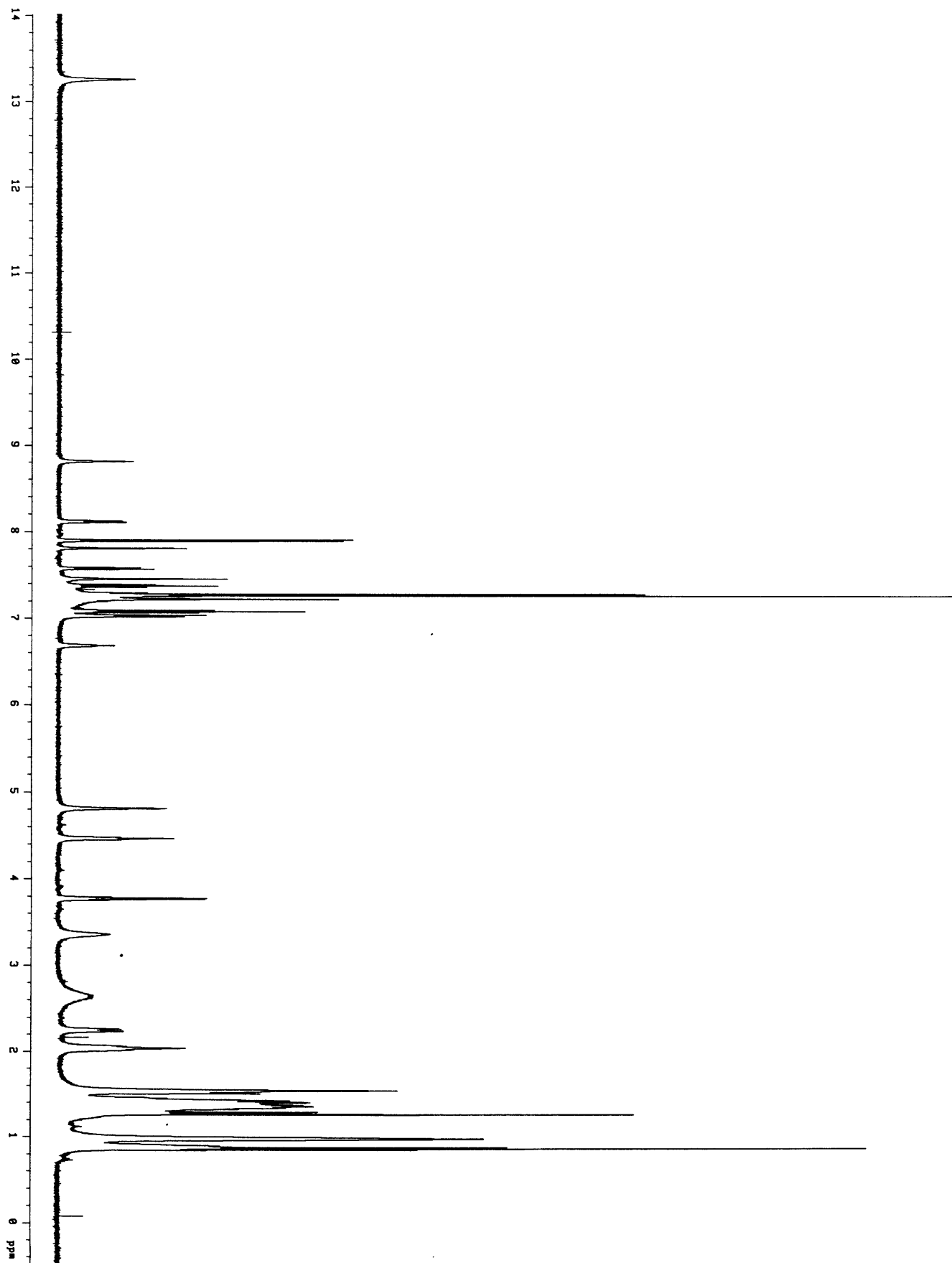


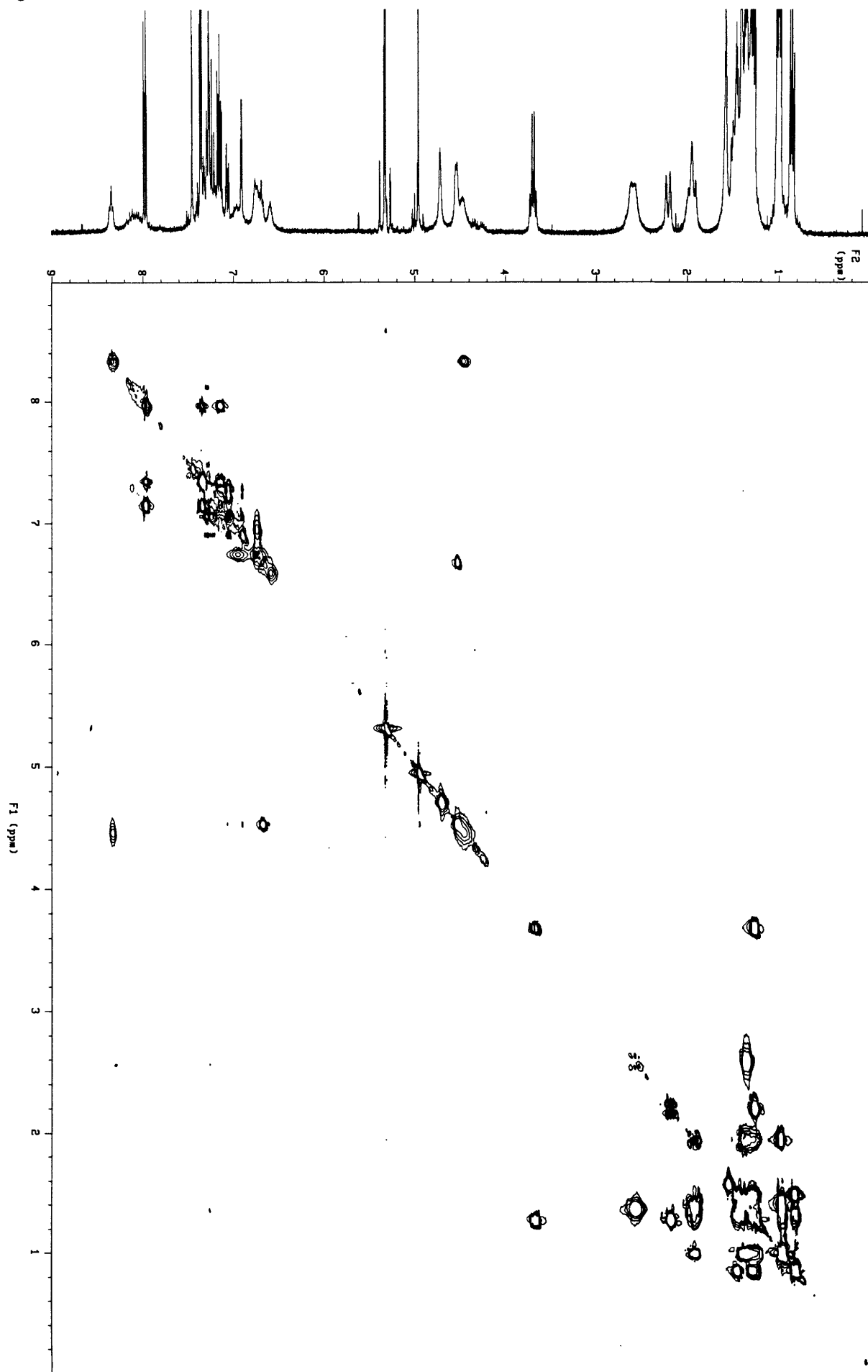


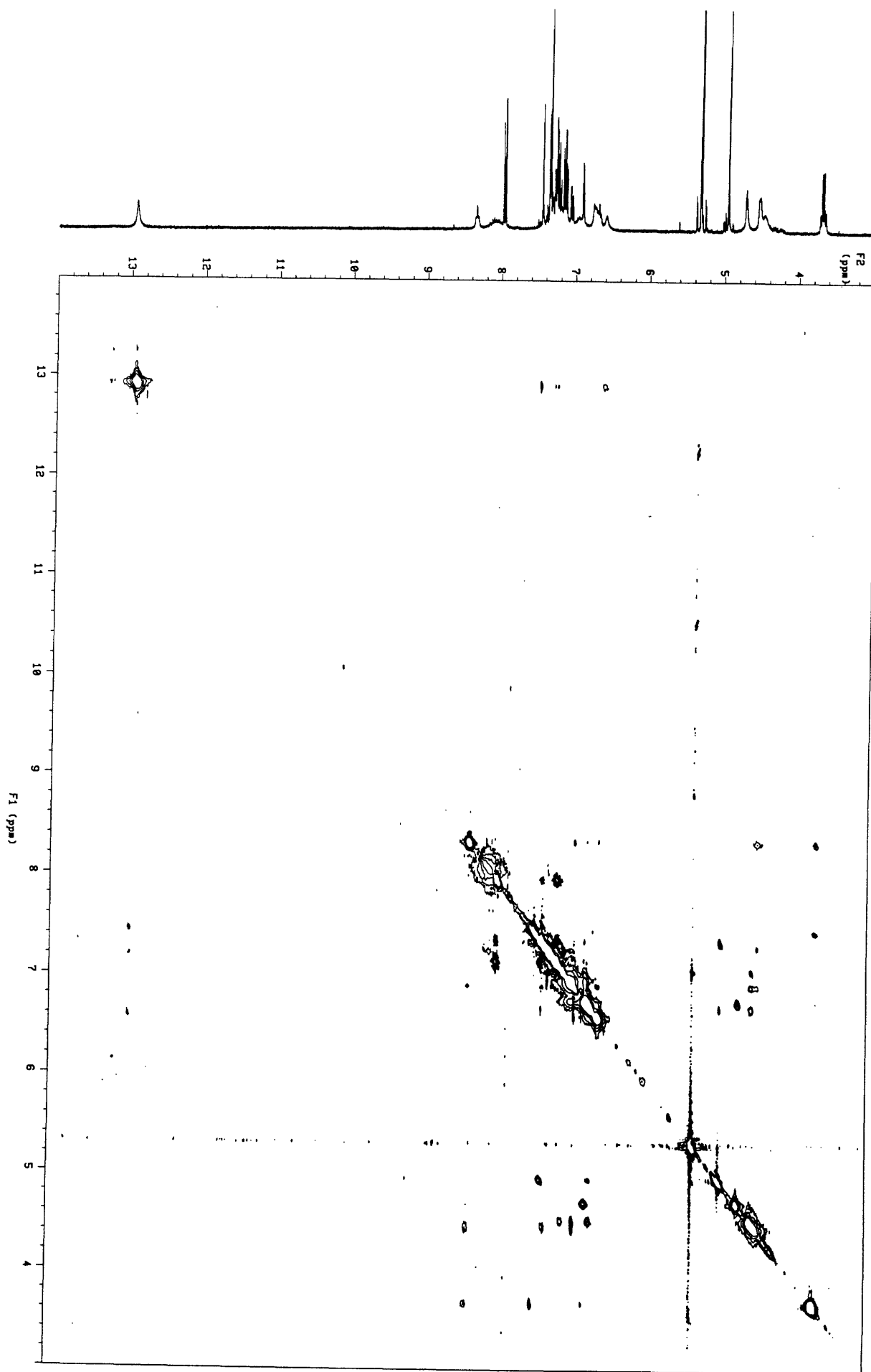


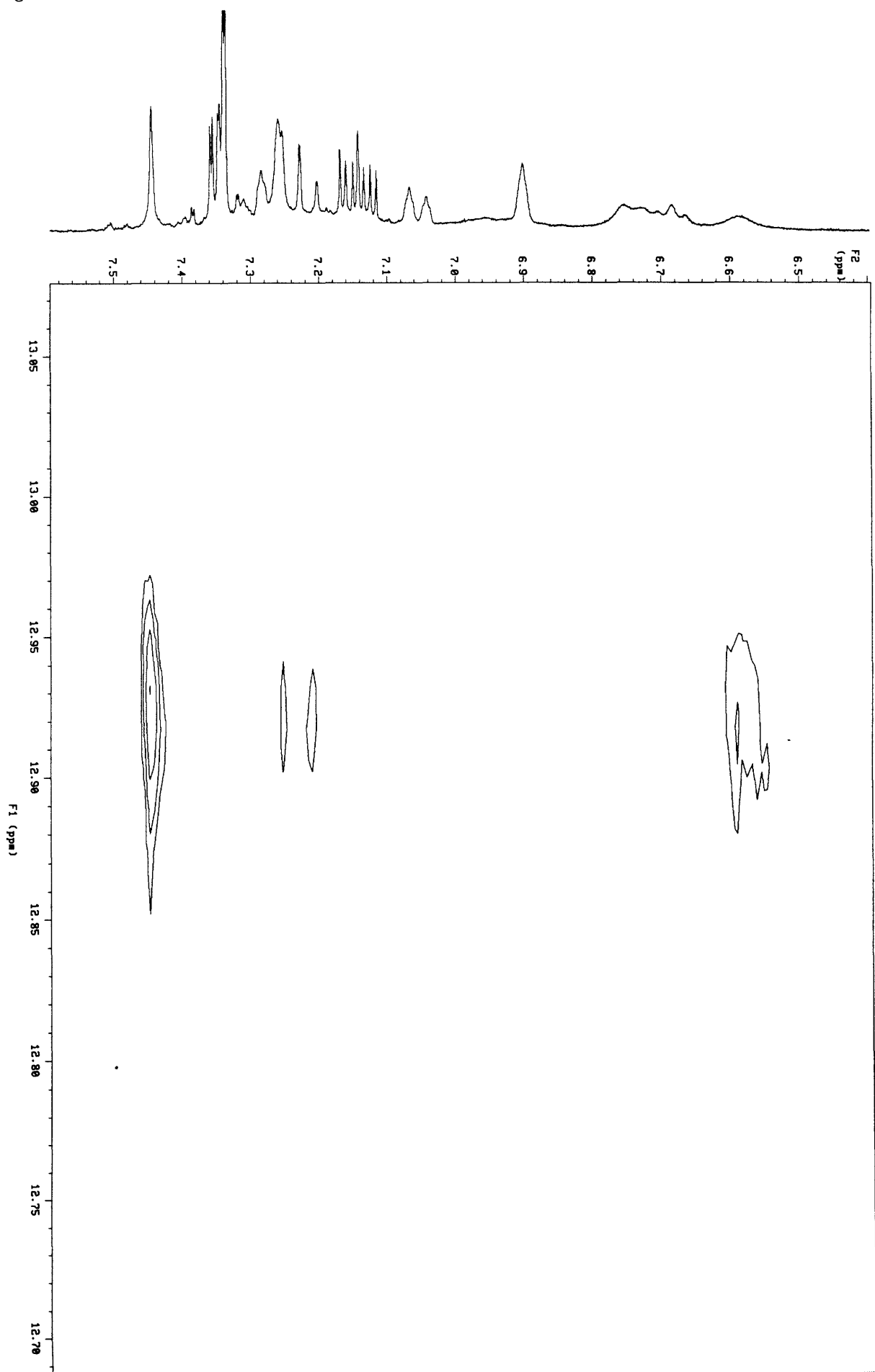












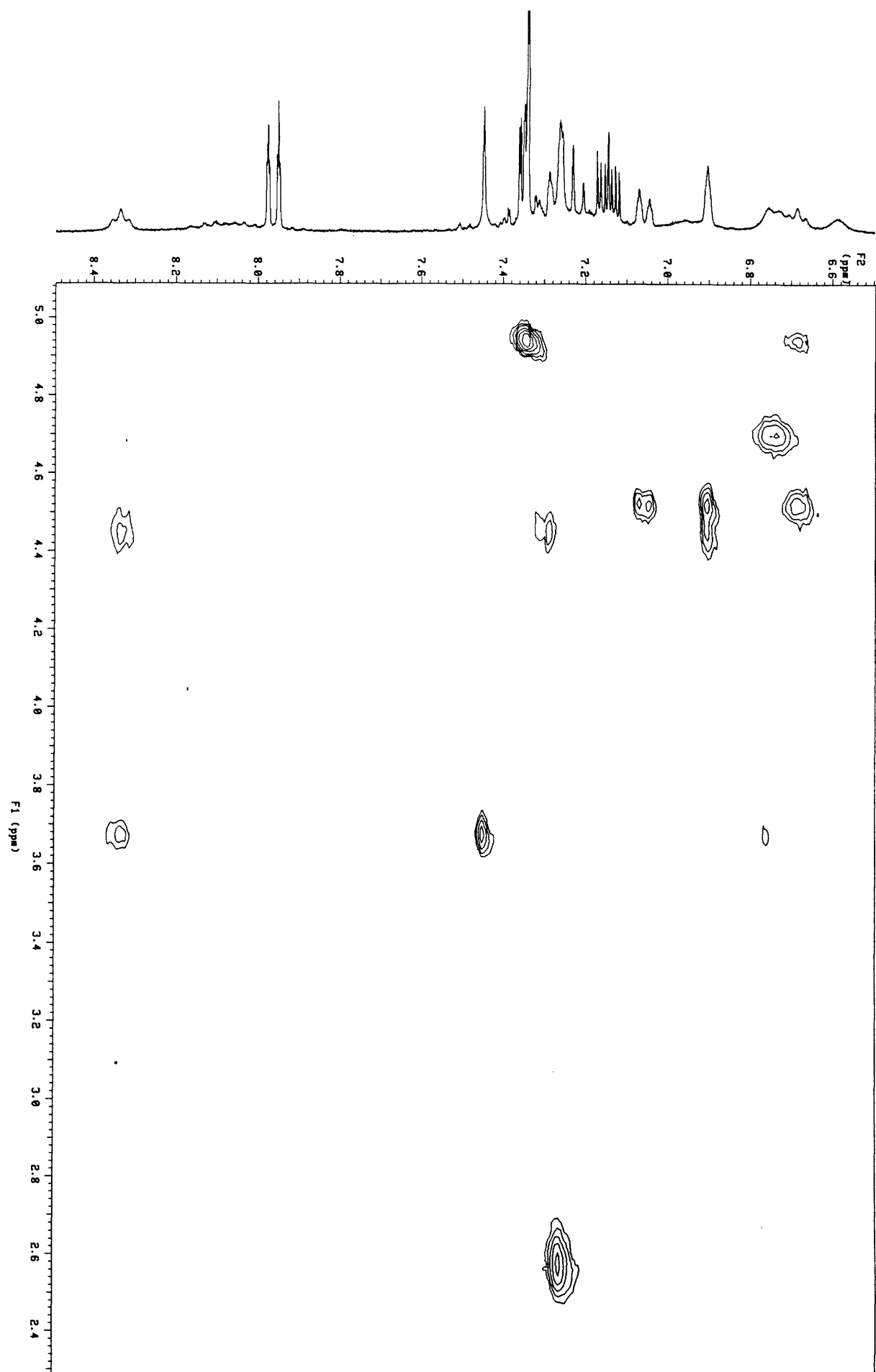


Figure E.5

