DESIGN AND CONFORMATIONAL ANALYSIS OF STRUCTURALLY-RELATED ANALOGUES OF POLYKETIDE NATURAL PRODUCTS

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

Erik M. Larsen

______________________________
Richard E. Taylor, Director

Graduate Program in Chemistry and Biochemistry
Notre Dame, Indiana
October 2016
DESIGN AND CONFORMATIONAL ANALYSIS OF STRUCTURALLY-RELATED ANALOGUES OF POLYKETIDE NATURAL PRODUCTS

Abstract

by

Erik M. Larsen

The conformation of a small molecule is crucial to its biological activity, as ideal binding requires both ligand and protein to spatially complement one another. This is particularly important when considering compounds such as the polyketide class of bioactive natural products, in which evolution has included a wide variety of structural features designed to limit local conformational preferences. As such, these molecules, particularly the macrocyclic varieties, will adopt preferred low-energy conformations. We believe that understanding these preferences for a given polyketide is a necessary precursor to analogue design, as any modifications to the structure should be mindful of how conformation will be affected. As such, our research uses techniques based around high-field NMR and molecular modeling experiments to analyze polyketide conformational preferences.

Neopeltolide is a 14-membered marine polyketide natural product that has demonstrated considerable antifungal and anticancer potential. Our conformational analysis of its macrolide found that it adopts a single conformation in solution, which minimizes syn-pentane interactions between substituents at the C9, C11, and C13
positions. Further analyses found that C2 substitution should not affect this conformation, and a methyl group was attached synthetically to probe any potential biological effects. Additionally, a preliminary analysis was made of neopeltolide’s purported biological target, cytochrome $bc_1$, for potential binding sites, which found a potentially new allosteric region that fits neopeltolide better than either of the known sites of inhibition.

Zampanolide and dactylolide are related 20-membered marine polyketides that exhibit considerable cytotoxicity against multiple human cancer cell lines and act as a microtubule stabilizing agents. Our research used material from a synthesis of dactylolide to elucidate its solution structure and conformational preferences. This work found that the shared macrolide interconverts between three major families in DMSO-$d_6$, one of which bears considerable resemblance to the bound conformation of zampanolide. We used this work to design a simplified analogue that omits the C17 methyl and C13 exocyclic alkene, which possessed a considerably shortened synthetic route in comparison with the parent compound.

Finally, the misidentification of the bioactive conformation of the polyketide epothilone prompted us to investigate whether the notable structural elements that influence polyketide conformation could be affected by a solid state environment. This research involved searching the Cambridge Structural Database for a variety of features, including $syn$-pentane interactions, the gauche effect, allylic strain and others. Our results found that high-energy features relying on steric effects will be unaffected by crystal packing forces, while features reliant upon lower-energy stereoelectronic effects or hydrogen bonds were far more likely to be influenced.
To my parents, for getting me this far.
## CONTENTS

Figures........................................................................................................................................... vii

Schemes ........................................................................................................................................... xv

Tables ............................................................................................................................................ xvii

Abbreviations ........................................................................................................................... xviii

Acknowledgments................................................................................................................... xxii

Chapter 1: Conformation-Activity Relationships of Polyketide Natural Products......... 1

1.1 Purpose....................................................................................................................................... 1

1.2 From Sachse to Barton: A Brief History of Conformational Analysis .................... 1

1.3 The Role of Conformational Entropy in Polyketide Molecular Recognition ....... 5

1.4 Conformational Control Elements in Polyketides ..................................................... 7

1.5 Conformational Mimics of Polyketide Natural Products ........................................ 8

1.5.1 Bryostatin.............................................................. 9

1.5.2 Laulimalide ......................................................... 15

1.5.3 Exiguolide .......................................................... 20

1.5.4 Discodermolide ................................................... 24

1.5.5 Dictyostatin ......................................................... 28

1.5.6 Trienomycin A .................................................... 32

1.5.7 Spongistatin .......................................................... 35

1.6 Conformationally-Restricted Analogues of Polyketide Natural Products ........... 40

1.6.1 Epothilone ........................................................... 41

1.6.2 Geldanamycin ..................................................... 47

1.7 Conclusions and Future Outlook ........................................................................... 52

Chapter 2: From NAMFIS to DISCON: Tools for Conformational Analysis ........... 53

2.1 Purpose....................................................................................................................................... 53

2.2 Tools for Analyzing Small Molecule Shape ................................................................. 53

2.2.1 Crystallography ...................................................... 54

2.2.2 NMR ................................................................ 55

2.2.3 Molecular Modeling ............................................. 56

2.3 NAMFIS Analysis of Molecular Models ................................................................. 59

2.3.1 Laulimalide .......................................................... 60

2.4 Polar Map Analysis of Molecular Models .............................................................. 63

2.4.1 Epothilone A .......................................................... 64
4.2.5.1 Taylor Lab Synthesis of Dactyloolide and Zampanolide . 151
4.3 Previous Conformational Work .........................................................................................154
  4.3.1 Porco: Conformation of the Hemiaminal Side Chain ..................................................154
  4.3.2 Diaz: Bioactive Conformation of Dactyloolide .............................................................157
4.4 Solution Conformational Analysis of Dactyloolide ...........................................................159
  4.4.1 Molecular Modeling of Dactyloolide ..............................................................................160
  4.4.2 1D and 2D NMR Experiments .......................................................................................162
  4.4.3 DISCON Identification of Dactyloolide Solution Conformations ..................................162
  4.4.4 Conformational Analysis of the Northern Fragment ......................................................165
  4.4.5 Conformational Analysis of the Eastern Fragment .........................................................166
  4.4.6 Conformational Analysis of the Western Fragment .......................................................167
  4.4.7 Comparison of DISCON-derived Solution Conformations to Tubulin-Bound Zampanolide ................. 169
4.5 Design of a Simplified Analogue: 17-Desmethyl Dactyloolide .........................................171
  4.5.1 Previous Analogues: Altmann 2012 ...............................................................................171
  4.5.2 Simplification of the C16-C17 Trisubstituted Olefin ......................................................175
  4.5.3 Conformational Analysis of Simplified Dactyloolide Analogue .....................................176
  4.5.4 Synthesis of 17-desmethyl-13-desmethylene Dactyloolide ............................................180

Chapter 5: The Effects of Crystallographic Packing on Common Polyketide Structural Features ................................................................................................................183
5.1 Purpose ............................................................................................................................183
5.2 The Journey to the Bound Conformation of Epothilone ..................................................183
5.3 The Effects of Environment on Ligand Binding .................................................................186
  5.3.1 Ligand Binding Strain ....................................................................................................187
5.4 Preferred Solid-State Conformations for Common Polyketide Features ................................189
  5.4.1 Methodology ................................................................................................................190
  5.4.2 Structural Features Affecting One Bond .......................................................................190
    5.4.2.1 Allylic 1,3-Strain .....................................................................................................190
    5.4.2.2 1,2-Diols – The Gauche Effect .................................................................................193
    5.4.2.3 Esters/Lactones .......................................................................................................196
    5.4.2.4 α-Substituted Ketones ............................................................................................197
  5.4.3 Structural Features Affecting Two Bonds .....................................................................200
    5.4.3.1 Syn-Pentane Interactions ........................................................................................200
    5.4.3.2 1,3-Diols ................................................................................................................203
    5.4.3.3 β-Hydroxy Ketones ...............................................................................................206
5.5 Conclusions .....................................................................................................................207

Chapter 6: Conclusions and Future Directions ..................................................................................209
6.1 Neopeltolide ....................................................................................................................209
  6.1.1 Neopeltolide as a Fluorescent Probe .............................................................................210
  6.1.2 Re-Evaluation of Cytochrome bc1 .................................................................................211
6.2 Zampanolide ....................................................................................................................212
  6.2.1 Non-Covalent ZampanolideAnalogue ...........................................................................214

Chapter 7: Supporting Information ..............................................................................................216
7.1 General Methods .................................................................................................................216
FIGURES

Figure 1.1 – Compounds explored by Baeyer and his analogous description thereof. .................................................2

Figure 1.2 – Sachse’s paper models of cyclohexane. From Russell, *van’t Hoff-Le Bel Centennial*. ..................................................3

Figure 1.3 – Mohr’s models for cis- and trans-decalin. .................................................................................................4

Figure 1.4 – Some of the common features of polyketide natural products used to influence conformation. A) syn-pentane interactions; B) β-hydroxy ketones; C) 1,3-diols; D) A13-strain; E) Gauche effect; F) esters and lactones; G) α-substituted ketones. ....................................................................................................................8

Figure 1.5 – Bryostatin 1.1 and 12-myristate-13-acetate (PMA) 1.2. .................................................................10

Figure 1.6 – Overlay of acetal 1.3 (blue) and 1.1 (green) solution conformations with hydrogen-bonding network. .........................................................................................................................12

Figure 1.7 – Major solution conformer of salicylate analogue 1.4.................................................................13

Figure 1.8 – Keck’s analogues 1.5 and 1.6; Krische’s analogue 1.7. .................................................................14

Figure 1.9 – Laulimalide 1.8, isolaulimalide 1.9, and 11-desmethyl laulimalide 1.10. ................................16

Figure 1.10 – Overlay of the 1.8 (yellow) and 1.10 (green) low-energy solution conformers. ..................................................................................................................................................18

Figure 1.11 – Overlay of laulimalide major solution conformer (yellow) with tubulin-bound structure (purple). ........................................................................................................................................19

Figure 1.12 – The natural product (−)-exiguolide 1.11.................................................................20

Figure 1.13 – Exiguolide analogues developed by the Fuwa laboratory. .................................................................21

Figure 1.14 – Overlay of 1.11 (gray), 1.12 (orange), 1.13 (green) and 1.14 (blue).................................................23

Figure 1.15 – Discodermolide 1.15. .................................................................................................................24

Figure 1.16 – Discodermolide’s major solution conformation and conformational control elements. .........................................................................................26
Figure 1.17 – Conformationally-simplified analogues of discodermolide. The remainder of the structure is identical to 1.15 has been omitted for clarity. .........................27

Figure 1.18 – Comparison of dictyostatin 1.18 and discodermolide 1.15 structures........28

Figure 1.19 – Overlay of dictyostatin s-trans solution structure (blue) with TR-NOESY conformation (green) and solid-state structure of discodermolide (yellow).........29

Figure 1.20 – Dictyostatin-discodermolide hybrid molecules........................................31

Figure 1.21 – Trienomycins A-F. ..................................................................................33

Figure 1.22 – Overlay of trienomycin 1.21 (green) and monoene A 1.22 (yellow) SYBYL-generated structures. ........................................................................34

Figure 1.23 – Monoene A 1.22 and E 1.23 analogues..................................................35

Figure 1.24 – The spongistatin class of natural products.................................................36

Figure 1.25 – Retained ABEF conformational structure of spongistatin 1 (green).........38

Figure 1.26 – Simplified spongistatin analogues 1.26 and 1.27 developed by the Smith lab.................................................................39

Figure 1.27 – The epothilone class of natural products. .................................................41

Figure 1.28 – Conformational analysis of the C1-C8 and C11-C15 regions of 1.29 ......42

Figure 1.29 – Conformational analysis of the C11-C15 region within Taylor’s C14-substituted analogues 1.30 and 1.31. .................................................................43

Figure 1.30 – (E)-9,10-dehydroepothilone B 1.32 and its major solution conformer. ......45

Figure 1.31 – C3-deoxyepothilone 1.33 and 2,3-dehydroepothilone 1.34. ....................46

Figure 1.32 – Overlay of major solution conformer (green) and tubulin-bound structure (orange) of 1.28.................................................................................................47

Figure 1.33 – Geldanamycin 1.35 and its analogues..................................................48

Figure 1.34 – Overlay of solid state 1.42 (orange) with protein-bound 1.35 (gray)........50

Figure 2.1 – Approximate $^3J_{HH}$ coupling constants for a given dihedral angle between two protons, depending on heteroatom substitution of R1-4..................56

Figure 2.2 – Superposition of the six conformers from the “supine” family. Figure taken from Thepchatri et al, J. Am. Chem. Soc., 2005, 127, 12838. .........................62
Figure 2.3 – Epothilone A polar coordinate maps for C1-C9. Figure taken from Taylor, R.E.; Zajicek, J. J. Org. Chem. 1999, 64, 7224. ....................................................65

Figure 2.4 – Solution conformations of epothilone A in CD$_2$Cl$_2$ and selected chemical shifts, NOEs and calculated distances (s = Strong, m = Medium). Figures from Taylor, R.E.; Zajicek, J. J. Org. Chem. 1999, 64, 7224........................................66

Figure 2.5 – The polyketide natural product peloruside A. ..................................................67

Figure 2.6 – Overlay of peloruside major solution conformers A (blue) and B (green) ...68

Figure 2.7 – Polar maps of Peloruside A taken with the MM2* force field. .................69

Figure 2.8 – Spontaneous equilibration of pyran lactol to open keto-alcohol .............70

Figure 2.9 – Overlay of peloruside A tubulin-bound conformation (gray) with proposed solution conformation B (green) ..............................................................71

Figure 2.10 – Representative solution conformations of (+)-spongistatin 1 in DMSO. a) Twisted – 57% b) Flat – 13% c) Flat (CD ring flipped up) – 8% d) Flat (CD ring flipped down) – 4% e) Overlay of a-d with respect to the ABEF ring system. .....73

Figure 2.11 – Comparison of global minima ab initio energies for peloruside A in MM2*, MM3* and OPLS-2001. Energies are relative to MM3*. ......................76

Figure 2.12 – Clustering of Conformers and Clustering Statistics windows for a sample library input........................................................................................................78

Figure 2.13 – Sample polar coordinate map and analysis of a dactyloide torsional angle (see Section 4.4.1 for full analysis). Each green dot on the plot indicates a single conformer. ..............................................................................................................80

Figure 2.14 – Sample DISCON dendogram from http://discon.sourceforge.net. Columns of identical or nearly identical color throughout may be safely omitted in future runs. ........................................................................................................................83

Figure 3.1 – The structure of neopeltolide as originally reported by Wright and co-workers and the revised structure reported by the Scheidt and Panek laboratories. Key NOE interactions for 3.1a are highlighted in red. ........................................87

Figure 3.2 – Stereoisomers of the neopeltolide macrolide synthesized by the Panek laboratory. The oxazole-containing side chain (R) has been omitted for clarity...89

Figure 3.3 – Screening compounds used by Kozmin and co-workers in their mode of action study. Areas of change on leucascandrolide A are highlighted in orange. .90

Figure 3.4 – Some of the known inhibitors of cytochrome bc$_1$. The two forms of coenzyme Q10 (ubiquinol and ubiquinone) are boxed at the bottom. ...............92
Figure 3.5 – Scheidt’s analogues of neopeltolide. .....................................................................94

Figure 3.6 – Maier’s side chain analogues of neopeltolide, with the macrolide (R) removed for clarity. IC₅₀ values were obtained against A549 cells. ......................95

Figure 3.7 – Floreancig’s side chain analogues of neopeltolide, with the macrolide (R) removed for clarity. GI₅₀ values were measured against HCT116 cells. ...............96

Figure 3.8 – Fuwa’s side chain analogues, with IC₅₀ values measured against A549 cells. The 8,9-dehydro macrolide (X) is used in each of the analogues. .........................97

Figure 3.9 – Maier and Scheidt macrolide analogues of neopeltolide. The oxazole side chain (R) has been omitted for clarity. .................................................................98

Figure 3.10 – Floreancig macrolide analogues, based on 8,9-dehydro macrolide 3.31. The oxazole side chain (R) has been omitted for clarity. All analogues were tested against HCT116 cells, with neopeltolide (3.1b) possessing GI₅₀: 0.77 nM. .........99

Figure 3.11 – Fuwa’s neopeltolide analogues. The oxazole side chain (R) has been omitted for clarity. All IC₅₀ values were taken against A549 cells unless indicated otherwise. .............................................................................................................100

Figure 3.12 – Overall SAR of neopeltolide. From Fuwa et al, Chem. Eur. J. 2013, 19, 8100-8110. ...........................................................................................................101

Figure 3.13 – Preliminary modelling work showing the proposed conformations that would account for the observed H9-H11 (green) and H11-H13 (red) NOE crosspeaks in the correct stereochemical assignment of neopeltolide. .........................103

Figure 3.14 – Polar maps of the neopeltolide macrolide. ........................................................104

Figure 3.15 – Overlays of two conformational families showing the difference in ester conformation. The two conformers represented in each family differ by >1 kJ/mol (MM3*). ...............................................................................................................105

Figure 3.16 – Expected conformational results of C2 methyl substitution on the macrolide s-cis ester .........................................................................................................106

Figure 3.17 – Comparison of the (R)-2-methyl and (S)-2-methyl macrocyclic epimer polar maps..................................................................................................................107

Figure 3.18 – Comparison of different C2-substituted macrolide O-1-2-3 polar maps...108

Figure 3.19 – The two possible ester orientations of Conformer A for (R)-2-methyl neopeltolide ..................................................................................................................109

Figure 3.20 – Backbone polar maps and global minimum conformation for 9-epi neopeltolide 3.35. .............................................................................................................110
Figure 3.21 – Backbone polar maps and global minimum conformation for 11-epi neopeltolide 3.29

Figure 3.22 – Backbone polar maps and global minimum conformation for 11,13-epi neopeltolide 3.1a.

Figure 3.23 – Backbone polar maps and global minimum conformation for 8,9-dihydroxy neopeltolide 3.34.

Figure 3.24 – Part of the ROESY spectrum for the neopeltolide macrolide illustrating several major cross-peaks.

Figure 3.25 – Selected NMR variables used in DISCON and the two expected major conformers. Conformer B was not significantly represented in the data.

Figure 3.26 – Cartoon representation of the Q-cycle. From Vennam et al, Chem. Bio. Chem. 2013, 14, 1745-1753.

Figure 3.27 – Biological data from Kozmin et al224 representing the inhibitory effects on cytochrome c oxidation (top left) and the difference absorption spectra.

Figure 3.28 – SiteMap results on cytochrome b subunit illustrating the two major binding sites, known inhibitors, and the location of the newly-discovered “Site 2”.

Figure 3.29 – Global SiteMap output for the 3H1I monomer (left) and the results of a “blind” docking to each of them (right).

Figure 3.30 – IFD fit of top scoring neopeltolide fit in “Site 2” and the corresponding Ligand Interaction diagram.

Figure 3.31 – IFD fit of top scoring neopeltolide fit in the Qo site and the corresponding Ligand Interaction diagram.

Figure 3.32 – Overlay of the frames showing side-chain movement in “Site 2” MD simulation after 20 ns, along with the kinetic/potential energy of the system (bottom left/middle) and RMSD experienced during the simulation (bottom right, blue = protein, red = neopeltolide).

Figure 3.33 – Snapshots from Qo/neopeltolide MD simulation illustrating Rieske iron-sulfur protein deviation on the neopeltolide-bound monomer (green) vs. the famoxadone-bound one (red). Also shown is the RMSD experienced by the system (blue = protein, red = neopeltolide).

Figure 3.34 – Graphs comparing Prime-MMGBSA binding energies to the natural log of analogue IC_{50} values. The “All Analogues” chart also includes the Fuwa side chain analogues illustrated in Figure 3.8.
Figure 3.35 – Comparison of possible C2 enolates. Major conformer (left), Z-enolate (middle), E-enolate (right). .................................................................138

Figure 3.36 – Possible explanation for the observed stereoselectivity of the C2 macrolide addition (below). Floreancig’s hydrogenation of 8,9-dehydro macrolide (above) is shown to illustrate a similar sterically-driven facial preference. .................139

Figure 4.1 – Zampanolide and dactyloolide. ....................................................................................................................144

Figure 4.2 – Reaction of methanol with dactyloolide.............................................................................................145

Figure 4.3 – Crystal structure of zampanolide bound to β-tubulin (RSC PDB: 4I4T). Hydrogen bonding interactions shown in green. ................................................148

Figure 4.4 – Off-resonance and STD spectra of dactyloolide taken in D₂O while bound to microtubules. The STD spectrum illustrates the protons affected by microtubule binding. ..................................................................................158

Figure 4.5 – TR-NOESY derived conformation of dactyloolide bound to tubulin. Figure taken from Field et al., Chem Biol., 2012, 19, 686. ..............................................159

Figure 4.6 – Polar maps of dactyloolide derived from combined MM3*, MMFF* and OPLS-2005 Monte Carlo conformational searches. .............................................161

Figure 4.7 – Dendogram from selected DISCON dactyloolide NMR parameters (flipped on side to fit page).................................................................163

Figure 4.8 – DISCON-derived representative conformations of 4.14. ............................................................164

Figure 4.9 – Conformational analysis of the C1-C7 region .................................................................166

Figure 4.10 – Conformational analysis of the C7-C11 enone region .................................................167

Figure 4.11 – Conformational analysis of the C15-C19 region .................................................................168

Figure 4.12 – Possible orientations and observed dihedral angle of the 17-18-19-O torsion. .................................................................169

Figure 4.13 – Overlay of major solution conformer A (green) with the tubulin-bound structure of zampanolide (orange) bound to H229 (gray). .........................170

Figure 4.14 – Side-chain analogues synthesized by the Altmann laboratory ........................................173

Figure 4.15 – Simplified macrocyclic analogues synthesized by the Altmann laboratory showing changes to the pyran region. .................................................................175

Figure 4.16 – Simplified analogues targeted for synthesis (4.26) and conformational analysis (4.27). ........................................................................................................................................177
Figure 4.17 – Low-energy conformations observed from a Monte Carlo conformational search of 4.27......................................................................................................178

Figure 4.18 – Observed conformational preferences of conformers D and E in the changed polar maps ........................................................................................................................................179

Figure 5.1 – EC-derived tubulin-bound conformation with observed syn-pentane interactions (PDB 1TVK, top left) compared to the solution/X-ray structure of epothilone A (top right), the CytP450K bound conformation (bottom left), and the TR-NOESY structure in unassembled tubulin (bottom right). ........................................184

Figure 5.2 – Epothilone analogues designed to be conformationally-restricted to the Nettles et al EC structure. .................................................................................................186

Figure 5.3 – Sample of Perola’s analysis applied to PDB 1EZQ. Global minimum conformation (left) is ~100 kJ/mol lower in MMFF potential energy than bound conformation (right)..................................................................................................................188

Figure 5.4 – Energy changes arising as a result of torsional changes in 5.4 and 5.5 ..........191

Figure 5.5 – Polar histogram for the dihedral angle θ representative of A 1,3-strain. ....192

Figure 5.6 – The gauche effect in 1,2-difluoroethane.................................................................................................193

Figure 5.7 – Crystal packing in BAGDOC (ammonium D-gluconate) demonstrating intermolecular hydrogen bonding. Both anti and gauche diols can be observed. 194

Figure 5.8 – Observed torsional angles for vicinal diols in non-macrocycles (left) and macrocycles (right). ..................................................................................................................195

Figure 5.9 – Observed torsional angles for OH-OMe gauche effect. .........................196

Figure 5.10 – Torsional preferences for esters and lactones.............................................197

Figure 5.11 – Preferred conformations of 2-butanone (left) and 3-methyl-2-butanone (right). ..................................................................................................................198

Figure 5.12 – Observed torsional angles for α-methyl ketones. Expected “skew” conformation (top right) vs. major observed anti conformation (bottom right). 199

Figure 5.13 – Observed O-C-C-O torsional angles for α-hydroxy ketones.................200

Figure 5.14 – Ramachandran plot for observed torsional angles of cis-1,3-dimethyl moieties..................................................................................................................201

Figure 5.15 – Ramachandran plot for observed torsional angles of trans-1,3-dimethyl moieties ..................................................................................................................202
Figure 5.16 – Ramachandran plot for the observed torsional angles of \textit{cis}-1,3-diol moieties. ..................................................................................................................204

Figure 5.17 – Ramachandran plot for the observed torsional angles of \textit{trans}-1,3-diol moieties. ..................................................................................................................205

Figure 5.18 – Ramachandran plot for observed torsional angles of \textit{β}-hydroxy ketone moieties. Ideal hydrogen-bonding regions are outlined in red. .................................207

Figure 6.1 – Comparison of \textbf{3.4} to known Q\textsubscript{o}-site inhibitors famoxadone and myxothiazol.................................................................212

Figure 6.2 – Covalent and non-covalent dactylolide analogues. .................................214
Scheme 1.1 – Representative transacetalization ring closure for bryostatin core..............11

Scheme 3.1 – Conditions for Taylor’s synthesis of β-alkoxyacrylate 3.54 en route to the macrolide core of neopeltolide. a) BH$_3$·SMe$_2$; b) (OMe)MeNH$_2$Cl, iPrMgCl; c) DMP, 86% (over 3 steps); d) 3.47, BF$_3$·OEt$_2$; e) BOMCl, Hünig’s base, 55% (over 2 steps); f) SmI$_2$, PhCHO (dr > 20:1); g) Me$_3$OBF$_4$, Proton Sponge, 73% (over 3 steps); h) ICl; then Na$_2$S$_2$O$_3$, 71% (dr > 20:1); i) ethyl propiolate, PBu$_3$, 98%. .....................................................................................................................135

Scheme 3.2 – Conditions for completion of the neopeltolide macrolide 3.57. a) AIBN, nBu$_3$SnH; 95 % (dr 19:1); b) KOH, MeOH, 95%; c) TCBCl, Et$_3$N, DMAP, 87%; d) H$_2$, Pd/C, 100%........................................................................................................................................................................136

Scheme 3.3 – Synthesis of (R)-2-methyl macrolide 3.58. .........................................................138

Scheme 3.4 – Route to side chain 3.65 devised by Leighton and co-workers. 243 .............140

Scheme 3.5 – Completion of the synthesis of neopeltolide 3.1b and (R)-2-methyl neopeltolide 3.68. ........................................................................................................................................................................141

Scheme 4.1 – Taylor retrosynthesis of zampanolide. ...............................................................151

Scheme 4.2 – Synthetic route to pyran 4.5 for cross-coupling .............................................152

Scheme 4.3 – Route to (−)-dactylolide 4.14 and (−)-zampanolide.......................................153

Scheme 4.4 – Hemiaminal model substrates prepared by the Porco laboratory.................155

Scheme 4.5 – Representative hydrogen bonding networks in hemiaminal diastereomers of 4.15 ........................................................................................................................................................................156

Scheme 4.6 – Retrosynthesis of simplified analogue 4.26. .....................................................180

Scheme 4.7 – Initial (failed) plan to install trisubstituted olefin in dactylolide/zampanolide core. ..................................................................................................................................................181

Scheme 4.8 – Synthetic route to target intermediate 4.28. .....................................................182
Scheme 6.1 – Conversion and attachment of coumarin-based dye 6.1. a) LiOH, THF, H₂O; b) pent-4-en-1-ol, EDAC, DMF; c) CatBH, MeCONMe₂ (cat.); d) p-TolSO₂I, t-BuON=NO=t-Bu; e) NaHMDS (5 eq), then 6.3. ........................................211

Scheme 6.2 – Route to non-covalent analogue 6.9 and its differences from the standard zampanolide route. .................................................................................................................................215
TABLES

Table 7.1 Neopeltolide Macrolide $^1$H NMR Coupling Constants ........................................222

Table 7.2 Dactylolide ROESY-Derived Interproton Distances Used in DISCON Calculations..............................................................223

Table 7.3 Dactylolide $^1$H NMR Coupling Constants Used in DISCON Calculations ....224
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>absolute optical rotation</td>
</tr>
<tr>
<td>$^{13}$C</td>
<td>carbon 13 NMR</td>
</tr>
<tr>
<td>$^1$H</td>
<td>proton NMR</td>
</tr>
<tr>
<td>Å</td>
<td>angstrom</td>
</tr>
<tr>
<td>AIBN</td>
<td>azobisisobutyronitrile</td>
</tr>
<tr>
<td>Bn</td>
<td>benzyl</td>
</tr>
<tr>
<td>BOM</td>
<td>benzyloxyethylmethyl</td>
</tr>
<tr>
<td>c</td>
<td>concentration</td>
</tr>
<tr>
<td>cm</td>
<td>centimeter</td>
</tr>
<tr>
<td>CSA</td>
<td>camphorsulfonic acid</td>
</tr>
<tr>
<td>δ</td>
<td>chemical shift</td>
</tr>
<tr>
<td>d</td>
<td>doublet</td>
</tr>
<tr>
<td>DBU</td>
<td>1,8-diazabicyclo[5.4.0]undec-7-ene</td>
</tr>
<tr>
<td>DDQ</td>
<td>2,3-dicyano-5,6-dichloro-parabenzoquinone</td>
</tr>
<tr>
<td>DEAD</td>
<td>diethyl azodicarboxylate</td>
</tr>
<tr>
<td>DIBAL-H</td>
<td>diisobutylaluminum hydride</td>
</tr>
<tr>
<td>DIPEA</td>
<td>N,N-diisopropylethylamine</td>
</tr>
<tr>
<td>DIPT</td>
<td>diisopropyl tartrate</td>
</tr>
<tr>
<td>DMAP</td>
<td>4-dimethylaminopyridine</td>
</tr>
<tr>
<td>DMF</td>
<td>dimethylformamide</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Meaning</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
</tr>
<tr>
<td>DMP</td>
<td>Dess-Martin periodinane</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>DTBP</td>
<td>2,6-di-tert-butylpyridine</td>
</tr>
<tr>
<td>dr</td>
<td>diastereomeric ratio</td>
</tr>
<tr>
<td>E</td>
<td>entgegen</td>
</tr>
<tr>
<td>ee</td>
<td>enantiomeric excess</td>
</tr>
<tr>
<td>EDCI</td>
<td>1-ethyl-3-(3-dimethylaminopropyl)carbodiimide</td>
</tr>
<tr>
<td>EPHP</td>
<td>1-ethylpiperidine hypophosphite</td>
</tr>
<tr>
<td>Et&lt;sub&gt;2&lt;/sub&gt;O</td>
<td>diethyl ether</td>
</tr>
<tr>
<td>EtOAc</td>
<td>ethyl acetate</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>HMDS</td>
<td>hexamethyldisilazana</td>
</tr>
<tr>
<td>HMPA</td>
<td>hexamethylphosphoramide</td>
</tr>
<tr>
<td>HOBT</td>
<td>hydroxybenzotriazole</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>HRMS</td>
<td>high resolution mass spectrometry</td>
</tr>
<tr>
<td>HWE</td>
<td>Horner-Wadsworth-Emmons</td>
</tr>
<tr>
<td>Hz</td>
<td>Hertz</td>
</tr>
<tr>
<td>IC&lt;sub&gt;50&lt;/sub&gt;</td>
<td>half maximal inhibitory concentration</td>
</tr>
<tr>
<td>i-Pr</td>
<td>isopropyl</td>
</tr>
<tr>
<td>J</td>
<td>coupling constant</td>
</tr>
<tr>
<td>LAB</td>
<td>lithium amidotrihydroborate</td>
</tr>
<tr>
<td>LAH</td>
<td>lithium aluminum hydride</td>
</tr>
<tr>
<td>LDA</td>
<td>lithium diisopropylamide</td>
</tr>
</tbody>
</table>
M .............................................................. molar
m .............................................................. multiplet
Me ........................................................... methyl
MeOH ........................................................ methanol
mg .......................................................... milligram
µg .......................................................... microgram
MHz ....................................................... megahertz
mL ........................................................ milliliter
µM ........................................................ micromolar
mmol ...................................................... millimol
n-Bu ........................................................ normal butyl
n-Pr ........................................................ normal propyl
nM ......................................................... nanomolar
NMM ...................................................... N-methylmorpholine
NMO .................................................... N-methylmorpholine oxide
NMR ........................................................ nuclear magnetic resonance
p ............................................................. pentet
PCB .................................................... p-chlorobenzyl
PCC ...................................................... pyridinium chlorocromate
PDB ......................................................... Protein Data Bank
Ph .......................................................... phenyl
PKS ...................................................... polyketide synthase
PMB ...................................................... p-methoxybenzyl
ppm ........................................................ parts per million
PPTS ...................................................... pyridinium p-toluenesulfonate

xx
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>q</td>
<td>quartet</td>
</tr>
<tr>
<td>s</td>
<td>singlet</td>
</tr>
<tr>
<td>SAR</td>
<td>structure-activity relationship</td>
</tr>
<tr>
<td>t</td>
<td>triplet</td>
</tr>
<tr>
<td>TBAF</td>
<td>tetra-$n$-butylammonium fluoride</td>
</tr>
<tr>
<td>TBDPS</td>
<td>tert-butyldiphenylsilyl</td>
</tr>
<tr>
<td>TBHP</td>
<td>tert-butyl hydrogen peroxide</td>
</tr>
<tr>
<td>TBS</td>
<td>tert-butyldimethylsilyl</td>
</tr>
<tr>
<td>TEA</td>
<td>triethylamine</td>
</tr>
<tr>
<td>TES</td>
<td>triethylsilyl</td>
</tr>
<tr>
<td>THF</td>
<td>tetrahydrofuran</td>
</tr>
<tr>
<td>TLC</td>
<td>thin layer chromatography</td>
</tr>
<tr>
<td>TMS</td>
<td>trimethylsilyl</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>Z</td>
<td>zusamen</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I was never supposed to write this. In all honesty, attending graduate school was primarily a decision of necessity rather than desire, thrust upon me by the realities of a recession and the inherent weaknesses of a bachelor’s degree in a world where they’ve become commonplace. However, the last six years have wrought considerable changes on me, not only from the knowledge I’ve acquired but the incredible people I’ve met and the relationships with them that I’ve been fortunate enough to gain. While this work, as with all research, was only made possible by standing upon the shoulders of giants, their accolades can wait until the end credits. We rarely have a chance to thank those who help us keep our balance while we do so, and while many of them will never read this it would be remiss of me to leave them unacknowledged.

First, I have to thank my advisor, Rich Taylor. Few mentors would allow their students the space to grow as he did with me, and fewer still would let them venture so far off the beaten path of their own expertise. His patience and unwavering support in the face of my many screw-ups and various attempts to find myself is more than I ever could have asked for, and certainly more than I deserved. He taught me that strong fundamentals are the bedrock of any chemist, and helped plant the seeds for my own academic aspirations. My committee deserves thanks in a similar vein; Seth Brown for teaching me that inorganic chemistry is more than a useless collection of metals and Olaf Wiest for teaching me that the world of chemistry is strange, wonderful, and cannot be
trusted to make sense under any circumstances. I also have to thank Kathy Peterson and Sarah West for giving me incredible freedom as a teaching assistant, which allowed me to experiment with teaching styles and come into my own as an educator. Finally, I owe no small amount of gratitude to Jim Parise for being my teaching mentor over the past few years. The opportunities he provided me helped me remember why I enjoyed being a tutor in my youth, and focused my meandering career path squarely into that of academia.

The Taylor group at Notre Dame has provided some of the most profound personal and professional experiences of my life. Virtually every member has left me with important knowledge, though space will permit me only to single out the most extraordinary of them. Joe Arico served as my first lab mentor, and was incredibly patient and helpful in getting my lab skills back to mostly respectable while providing the foundations of the neopeltolide project. Cole Stevens provided great conversations, great advice, and great 4th of July parties, as well as being one of the very few people on this Earth that I can trust to always be honest with me whether I like it or not. Chia Fu Chang was a joy to work with both in lab and on the basketball court, despite the bruises he would respectively leave on my ego and my body. He also was kind enough to take the time out of his insane schedule to help me finish my experimental work. Finally, both Ian Harrier and Jarred Pickering have become the kind of friends I thought I’d never have, putting up with immeasurable amounts of my crap while providing invaluable advice and a sympathetic ear whenever I needed it. I can only hope that I’ve been as good a friend in return.

Two of my colleagues in particular need to be singled out for making this manuscript even possible. Matt Wilson provided the synthetic material that made the
conformational analysis of dactylolide and zampanolide possible, and contributed greatly to the conformational review of polyketides. The work we performed together would not have been possible without his skills and his enthusiasm. In all likelihood I would not have remained in graduate school had he not been generous enough to offer me the chance to work with him, and been kind enough to keep me motivated when I was out of ideas. In the same vein, I have to thank Marcus Arieno for taking an unbelievable amount of time to personally guide me through docking and molecular dynamics, and untold more hours answering my unending questions. Beyond that, the time we spent outside of lab was equally valuable in keeping me sane and grounded as the years in graduate school dragged on, and I’ll miss our weekly breakfasts dearly.

Several people outside of the Notre Dame bubble were also instrumental in getting me this far. From Alma College, I would like to thank my advisor Scott Hill for first introducing me to the wonders of organic chemistry, and for recommending me for graduate school despite my less-than-sterling undergraduate career. From MPI Research, I have to thank Ed Magnuson for managing to instill something resembling a work ethic in me during our short time together. Above all, I have to thank my family, my brother Brian and my parents Scott and Martha, for reasons that if listed here would dwarf the remainder of this manuscript. Suffice to say that their love, their trust, and their unwavering belief in me even when I have none deserve a debt of gratitude that goes far beyond my ability to repay.

Finally, thank you, dear reader, for taking the time to pull this tract from its dusty server, even if only for a few minutes. Though I may have frequently wished this manuscript to fall into a black hole during its creation, I appreciate that you thought it
might be useful. Even if you put it down having gleaned only that it contained the ramblings of an idiot, I hope that this section has reminded you of the people that make your own research possible. Go thank them once you’re done here.
1.1 Purpose

This chapter will cover how conformational analysis has been used to guide the design of biologically-active analogues of polyketide natural products through the establishment of conformation-activity relationship. A brief history of conformational analysis in organic molecules will be outlined, followed by a discussion of conformational entropy in small molecule binding. Conformational studies on polyketide natural products that have led to analogue design will be discussed, including ones that focus on mimicking the conformational profile and ones that seek to rigidify otherwise flexible structures to explore the bioactive conformation.

1.2 From Sachse to Barton: A Brief History of Conformational Analysis

Although Derek Barton is generally credited with launching the field of conformational analysis of organic molecules, the history behind the three-dimensional structure of carbon rings stretches back considerably further. In the course of his work with substituted cyclohexanes, Adolf von Baeyer reported that hexahydromellitic acid and hexahydroterephthalic acid (1,2,3,4,5,6- and 1,4-cyclohexanecarboxylic acid,
respectively) each possessed two distinct forms. Baeyer explained these observations in terms of planar rings, using olefins as an analogy (Figure 1.1).

![Hexahydromellitic acid](image1)

hexahydromellitic acid

![Hexahydrotartratic acid](image2)

hexahydrotartratic acid

Figure 1.1 – Compounds explored by Baeyer and his analogous description thereof.

However, the spontaneous interconversion between the two forms piqued the curiosity of an assistant at the Charlottenburg (now Berlin) Institute of Technology named Hermann Sachse. Sachse disagreed with Baeyer’s assertion that cyclohexane rings were slightly strained by nature, postulating two unstrained forms of cyclohexane based on the tetrahedral geometry of carbon that he called “symmetrical” and “unsymmetrical”. His initial publication provided instructions on building models out of cardboard that would illustrate these two forms, now recognizable as the familiar chair (symmetrical) and boat (unsymmetrical) conformations, and suggested that they could interconvert and flip under certain conditions to provide distinct forms (Figure 1.2). However, because Sachse presented his work in the form of mathematical language and
trigonometric proofs, his arguments could not be understood by most chemists of the
time. Despite several attempts at publishing and explaining his ideas, his disagreement
with the position of someone with the clout of Baeyer failed to win him much support,
and he died in obscurity in 1893 at the age of 31.

Figure 1.2 – Sachse’s paper models of cyclohexane. From Russell,
*van’t Hoff-Le Bel Centennial.*

While Sachse’s ideas were not viewed unfavorably, they were largely ignored by
the scientific community. Although chemists realized that unstrained conformations made
sense from a geometric perspective, they were unable to find experimental verification.
As a result, the consensus soon arose that even if ring-flip of a non-planar conformation
did occur, the result would still be a planar average. It was not until 1915 that Ernst Mohr
attempted to rescue Sachse’s theory. Mohr not only clearly illustrated his point by using
ball-and-stick figures instead of Van’t Hoff tetrahedra, but also asserted that cyclohexane
ring-flip, which was thought to require strong heat, could take place at room temperature.\textsuperscript{4,5} Although this lent credence to the planar average theory, Mohr recognized that ring-flip could not occur if the ring in question was locked as part of a larger molecule. As evidence, he pointed to Bragg’s recently-elucidated crystal structure of diamond, in which the chair structure was a clearly visible and repeating motif.\textsuperscript{6} Mohr also proposed that this ring-locking effect would be present in a structure as simple as decahydronaphthalene (decalin), which should exist in two forms that could not interconvert without the assistance of a broken chemical bond. W. Hückel proved this locking effect in 1925, separating decalin into two isomers via fractional distillation that he termed \textit{cis}- and \textit{trans}-decalin.\textsuperscript{7}

![Figure 1.3 – Mohr’s models for \textit{cis}- and \textit{trans}-decalin.](image)

However, the dual-boat model of \textit{cis}-decalin suggested by Mohr’s model and Hückel’s results was considered unlikely to Odd Hassel, who proposed that the significant number of eclipsed and flagpole-flagpole interactions between hydrogens within the structure would be sterically unfavorable.\textsuperscript{8} He eventually proved via electron diffraction that the \textit{cis} ring fusion of decalin was the result of a different orientation of two chairs,\textsuperscript{9} but his most significant contribution to the fledgling field of conformational
analysis was the distinction between axial and equatorial cyclohexane bonds. Hassel recognized that ring-flip would convert an axial substituent into an equatorial one (and vice versa). Even more importantly, the equatorial-substituted derivative possessed a lower energy than the axial one, providing the missing piece of information that explained why many of Sachse’s contemporaries could not verify his theories via normal chemical separation. In 1950, Barton established the utility of conformational analysis with his landmark paper on the conformation of the steroid nucleus, marking the final transition of the field from theoretical to practical.

1.3 The Role of Conformational Entropy in Polyketide Molecular Recognition

A small molecule’s biological activity is dependent on the interactions it makes with a target receptor’s binding site. Optimum binding requires a complementary spatial relationship of structural features to achieve favorable ligand-protein interactions. Synthetically-derived compounds in medicinal chemistry are frequently rigid structures, which limit the number of rotatable bonds and have become associated with “drug-like” properties. In contrast, bioactive natural products often possess significant degrees of conformational flexibility that may impart potentially beneficial properties for transport, solubility, selectivity, and binding. Nevertheless, these apparently flexible molecules will typically adopt preferred conformational profiles. In the case of the polyketide class of natural products, we can see many of the structural features that evolution includes to impart conformational preferences. Some of the hallmarks of this class include sp³-stereogenic centers (methyl and hydroxyl substituents), sp²-hybridized carbons (E- and Z-alkenes and carbonyl groups) and single bonds with restricted rotation (s-sp²-sp²) such as amides, esters, and polyenes. Even when uninvolved in binding, the steric and electronic
interactions these functionalities introduce serve to limit the number of potential low-energy conformations, while maintaining overall backbone flexibility. As such, these highly-substituted molecules preferentially adopt certain conformations, giving us what Hoffmann aptly described as “flexible molecules with a defined shape.” As long as the profile contains a conformer population complementary to the active site, the resulting change in free energy upon binding becomes more favorable.

The role of conformation in ligand-protein binding reaches beyond the bound state. Binding affinity involves both enthalpic and entropic contributions. While optimizing a drug candidate ideally involves the improvement of both terms, this is rarely achievable in practice. Additionally, improving enthalpic contributions (the direct non-covalent interactions between the ligand and the binding site) can introduce unforeseen complications. For instance, installing functional groups that increase ligand affinity for a target protein through enthalpic means will typically also affect water affinity. If the new groups form suboptimal protein contacts, the resulting effects on solvation and solubility can negatively influence both enthalpic and entropic parameters. As classic structure-activity relationships typically focus on the optimization of target-ligand interactions, they often ignore the potential impact these modifications have on ligand conformational preferences. Therefore, it is critical for one to consider conformational effects when applying a traditional medicinal chemistry approach to polyketide analogue design, evaluating both the structure-activity relationship and the conformation-activity relationship. Recent efforts from several labs have demonstrated the multipurpose nature of structural features within polyketide natural products and several examples are included below.
1.4 Conformational Control Elements in Polyketides

As researchers discovered in the mid-20\textsuperscript{th} century, conformation plays an important role in a variety of chemical transformations. The rate acceleration observed in anti-periplanar elimination reactions is one such example. Therefore, it comes as no surprise that many stereochemical control elements also impact polyketide conformation. As the attenuation of conformational freedom represents the ultimate goal in the evolutionary design of these biologically-active privileged structures, the most powerful restrictions utilize steric interactions between substituents or modifications of the carbon skeleton itself in the form of unsaturation.

One must take into account general steric minimization when considering a molecule’s preferred conformation, particularly when looking at cyclic molecules. The preference of substituents in six-membered rings to adopt equatorial rather than axial orientations to minimize gauche interactions is a fundamental example of this principle. However, the picture becomes more complicated as rings increase in size, flexibility, and heteroatom substitution. In rings larger than six members, strain arises as bond angles deviate from their idealized values and additional transannular interactions are introduced, which in concert with the appropriate substituents will force specific conformations over others. In large macrocycles with minimized ring strain, local acyclic structural features and subtle interactions between remote substituents will ultimately control the molecule’s conformational preferences, such as the ones seen in Figure 1.4. More details on some of these features will be covered in Chapter 5.
Figure 1.4 – Some of the common features of polyketide natural products used to influence conformation. A) syn-pentane interactions; B) β-hydroxy ketones; C) 1,3-diols; D) A\(^1,3\)-strain; E) Gauche effect; F) esters and lactones; G) α-substituted ketones.

1.5 Conformational Mimics of Polyketide Natural Products

Polyketide natural products contain a variety of diverse and synthetically-demanding functionalities, typically possessing extensive carbon skeletons with multiple stereogenic centers. Since these compounds usually exist in limited natural supply, investigators often rely upon total synthesis to provide sufficient material for therapeutic studies. Subsequently, the design of analogues that simplify or eliminate moieties with no effect on biological activity represents a worthwhile goal. Structural simplification not only shortens a molecule’s synthetic step count but also increases material throughput for clinical investigations. The Wender laboratory has termed this approach function-oriented synthesis,\(^15\) wherein biologically-active lead structures are simplified to incorporate only the activity-determining features. From a conformational standpoint, this design strategy involves identifying structural elements that minimally affect either the conformation of
regions responsible for protein interactions or the overall ensemble, ultimately creating a simplified structural analogue which retains the activity of the parent natural product.

1.5.1 Bryostatin

Bryostatin’s recent preclinical success exemplifies the therapeutic importance and potential impact of polyketide research on human health. Isolated by Pettit and co-workers from the marine bryozoan Bugula neritina, this diverse collection of polyketides contains over 20 complex natural products. The bryostatin family shares many important structural features such as three fully-functionalized tetrahydropyran rings, a unique methoxycarbonyl methylidene group, and several sites of oxygenation (Figure 1.5). These small molecules elicit a wide array of biological responses such as restoring apoptotic function in cancer cells, improving memory in animal models, and inducing latent HIV activation. Bryostatin’s impressive range of effects has been attributed to its ability to activate protein kinase C (PKC) by binding to the C1-domain. Interestingly, bryostatin shares many of these properties with phorbol esters such as PMA despite being structurally dissimilar, even competing with it for binding. However, bryostatin 1 lacks several of PMA’s unwanted biological responses, in particular its tumor-promoting capabilities, and will even antagonize them when dosed simultaneously.
Unfortunately, bryostatin’s scarce supply has hindered its further advancement into clinical trials for the treatment of cancer and Alzheimer’s disease. Although synthetic work over the years by numerous laboratories significantly improved the overall step count, these elegant syntheses have produced limited quantities of material.\textsuperscript{23} Additionally, isolation from \textit{Bugula neritina} yielded bryostatin in extremely limited amounts (approximately 18 g were isolated from 40,000 L of wet bryozoan),\textsuperscript{37} making natural extraction methods environmentally costly. Early computational and structure-activity relationship studies revealed that alterations within the C4-C16 region minimally affected bryostatin’s cytotoxicity profile.\textsuperscript{38,39} In contrast, deletion or alteration of functionality within the C15-C34 region produced analogues with reduced PKC affinity. Based on these findings, the Wender group hypothesized that exchanging the stereochemically complex C4-C16 region with a tetrahydropyran spacer would yield a simplified scaffold, enabling access to larger quantities of material for further biological
studies. To this end, the Wender group designed and synthesized a variety of simplified analogues based around an efficient macrocyclic acetalization reaction (Scheme 1.1).  

Scheme 1.1 – Representative transacetalization ring closure for bryostatin core.

From their new pool of analogues, the Wender laboratory selected 1.3 for detailed NMR analysis as a means of gauging how well its solution structure would match the predicted conformation. High-field 1D- and 2D-proton NMR experiments in benzene-$d_6$ yielded spectral data with coupling constants that significantly deviated from rotationally-averaged values, suggesting that 1.3 exists predominantly as a single conformation at room temperature. While the interproton distances failed to match the computationally-derived conformation calculated previously for bryostatin 10, a constrained gas-phase molecular dynamics simulation found several conformations within 2 kcal/mol of the global minimum. These low-energy conformers satisfied the distance constraints and matched nicely with both reported crystal and solution structures of bryostatin despite the
removal of all A- and B-ring substituents and the conversion of the C14 carbon to oxygen. Wender and co-workers attributed this conformation to a transannular hydrogen-bonding network between the C3 and C19 hydroxyl groups and the B-ring acetal oxygen (Figure 1.6).

![Figure 1.6 – Overlay of acetal 1.3 (blue) and 1.1 (green) solution conformations with hydrogen-bonding network.](image)

Biological analysis revealed that 1.3 and several other structurally related analogues successfully bound PKC isozymes in an established bryostatin assay, with the compound exhibiting nanomolar affinity consistent with multiple bryostatins. Furthermore, 1.3 displayed potent activity against several human cancer cell lines. This data further supported their hypothesis that bryostatin binding requires two functional domains: a recognition domain that interacts with the receptor and a spacing domain that properly orients and constrains the former. The Wender group has since synthesized a
multitude of related compounds investigating further modifications to the spacer domain, creating a library of bryostatin analogues which display single-digit nanomolar PKC affinity.\textsuperscript{22}

![Diagram of a molecule]

**Figure 1.7 – Major solution conformer of salicylate analogue 1.4.**

Recently, the Wender group began development of second-generation analogues that replace the spacer domain with a salicylate subunit, a commercially available motif that allows the macrocycle to retain key portions of the internal hydrogen-bonding network (Figure 1.7).\textsuperscript{43} Computational analysis suggested that salicylate 1.4 maintained a proper spatial orientation within the recognition domain wherein the C3-ether could make a hydrogen bond with the C19-hemiketal and the aromatic ring would associate with the membrane. Wender and co-workers synthesized 1.4 in only 23 total steps by employing a highly convergent and step-economical strategy. Biological evaluation showed that this
compound bound PKC with low nanomolar potency, approaching that of 1.1. Given that this analogue required even fewer steps than 1.3 and reduced the molecule’s structural complexity while retaining activity, the simplified salicylate substructure could be a promising lead scaffold for the development of future compounds.

Figure 1.8 – Keck’s analogues 1.5 and 1.6; Krische’s analogue 1.7.

In 2014, the Keck and Krische laboratories sought an explanation for the interesting biological responses provoked by bryostatin and related structural analogues with regard to the phorbol esters. During the process of designing simplified structures, the Keck group discovered that one of their newly-synthesized bryostatin analogues exhibited PMA-like biological responses. Subsequent analogues indicated that bryostatin-like activity depended upon the A-ring substituents but not substituents on the B-ring, suggesting that the B-ring could be safely removed altogether. The Keck and Krische laboratories independently synthesized two analogues lacking the B ring, 1.5 and 1.7 (Figure 1.8). Curiously, 1.5 was the minor product from the final deprotection step, with a ring-expanded analogue 1.6 representing the bulk of the material. Conformational and computational analysis of 1.5 and 1.7 alleged that both compounds retained the
hydrogen-bonding network and solution conformation of bryostatin, although the Keck laboratory also concluded that the ring strain arising from the C15-C17 unsaturated ester of 1.5 prompted the undesired ring expansion. Interestingly, 1.5, 1.6, and 1.7 all exhibited PMA-like biological responses in U937 histiocytic lymphoma cell assays, suggesting that the B-ring plays a pivotal role in bryostatin-like behavior. Recent work on neristatin 1, which binds to PKC and shares the A/B ring system of 1.1, has indicated that it also displays bryostatin-like activity. The bryostatins demonstrate how structural units can act as both conformational control and protein interaction elements, and how identifying the essential structures for control through conformational analysis can help efficiently design analogue scaffolds. The interesting reports from Keck and Krische suggest that bryostatin’s non-PMA mode of action depends on more than its ability to merely bind PKC, which may require more stringent evaluation of simplified bryostatin analogues moving forward.

1.5.2 Laulimalide

Laulimalide 1.8, a polyketide isolated from the sponge Cacospongia mycofijiensis, exhibits low nanomolar cytotoxicity against multiple cancers and retains significant activity against paclitaxel-resistant cell lines. The potent microtubule-stabilizing agent also binds tubulin at a non-taxane binding site and acts synergistically with paclitaxel and other taxane binders. In the interests of exploring laulimalide’s therapeutic potential, the Eisai Research Institute synthesized gram-scale quantities of 1.8 for in vivo biological testing. Though laulimalide exhibited a favorable pharmacokinetic profile compared to paclitaxel, it unfortunately displayed poor tumor growth inhibition accompanied by severe toxicity and mortality. These severe side effects precluded
dosing at higher concentrations. As such, improving laulimalide’s therapeutic usefulness requires analogue development.

Figure 1.9 – Laulimalide 1.8, isolaulimalide 1.9, and 11-desmethyl laulimalide 1.10.

Synthetic efforts spurred by laulimalide’s impressive in vitro activity revealed that 1.8 spontaneously degrades under mildly acidic conditions to yield isolaulimalide 1.9, a side product that displays significantly reduced cytotoxicity (Figure 1.9). Early analogue efforts circumvented this undesired degradation pathway primarily through removal/modification of the C16-C17 epoxide, the C2-C3 alkene, or capping the C20 alcohol.54,55 However, these relatively simple modifications diminished laulimalide’s cytotoxicity profile.56 Recognizing the importance of conformation to the analogue design process, several researchers investigated laulimalide’s conformational preferences to gain a better understanding of the molecule’s pharmacophore.

In 1996, Jefford and co-workers provided the first insights into laulimalide’s conformational preferences using X-ray crystallography.57 While this afforded valuable information on the molecule’s predominant conformation, the elucidation of laulimalide’s
bound conformation required a more rigorous analysis of the macrolide’s solution behavior. Nearly a decade later, the Paterson and Snyder groups analyzed the molecule’s solution conformation utilizing two different approaches. The Paterson group used an NMR-constrained conformational search\textsuperscript{58} while the Snyder group employed the NAMFIS (NMR Analysis of Molecular Flexibility in Solution) method.\textsuperscript{59}

Specifically, Paterson and co-workers used experimental coupling constant values for predicting dihedral angle preferences and NOESY experiments for identifying critical through-space interactions. They generated files with constrained torsional angles and distances and performed a restrained 10,000-step Monte Carlo conformational search with MacroModel and the MM2 force field, which provided several low energy conformers within 2 kcal/mol of the global minimum. While this analysis afforded low energy conformers resembling laulimalide’s solid-state conformation, the relative contribution of these solution conformers was left unexplored. The Snyder laboratory filled this gap in methodology by using NAMFIS to change the mole fraction of each conformer until a sum of square differences between the experimental NMR data and computed variables gave a “best fit” dataset. Their analysis and the NAMFIS methodology will be covered in greater detail in Section 2.3. Interestingly, both groups found that the primary solution conformers bore a great deal of similarity to laulimalide’s solid-state conformation. The solid-state crystal structure and major solution conformation showed that \textbf{1.8} preferred an open and gross-flattened conformation wherein the side chain, the C2-C3/C21-C22 unsaturated moieties, and the exocyclic epoxide all lie in the same plane. The polyketide’s dihydropyran rings also adopted a half-chair conformation and curved slightly under the macrocyclic ring.
Figure 1.10 – Overlay of the 1.8 (yellow) and 1.10 (green) low-energy solution conformers.

Given the difficulty with accessing laulimalide from isolation and total synthesis efforts, the Paterson and Wender groups designed a simplified analogue lacking the C11-Me, 11-desmethylauimalide 1.10 (Figure 1.9). Both laboratories targeted this particular functional group since its removal would decrease the synthetic step count with respect to those leading to 1.8. High-field NMR experiments and Monte Carlo conformational searches also saw a notable degree of flexibility within the C9-C12 region, which suggested that removing the C11-Me would not significantly impact laulimalide’s overall shape or biological activity (Figure 1.10). As a result, the Wender and Paterson groups synthesized the simplified analogue 10 utilizing highly convergent strategies. Notably, removing the C11-Me group reduced the Wender group’s synthetic step count and decreased the starting material cost by 22-fold. High-field 1D-NMR experiments showed that 10 retained similar $^3J_{H-H}$ coupling constants throughout the
northern and southern regions of the molecule. In particular, the 11-desmethyllaulimalide H14a/b-H15 dipolar couplings nearly matched identically with that of laulimalide \textbf{1.8}. However, the $J_{H10a-H11a}$ and $J_{H11-12a}$ coupling constants deviated slightly from laulimalide’s normal dipolar couplings indicating a minor conformational bias within the C9-C12 region.

![Diagram](image)

Figure 1.11 – Overlay of laulimalide major solution conformer (yellow) with tubulin-bound structure (purple).

Biological evaluation revealed that \textbf{1.10} possessed nanomolar cytotoxicity against multiple cancer cell lines including cell lines overexpressing the P-glycoprotein pump, though to a lesser degree than \textbf{1.8}.\textsuperscript{62} \textbf{1.10} also promoted tubulin polymerization at a critical concentration of only 1.1 ± 0.1 µM. Recently, Prota and co-workers published a co-crystal structure of laulimalide bound to tubulin that conclusively defines the molecule’s binding site on β-tubulin and fully elucidates the bioactive conformation.\textsuperscript{63} Interestingly, laulimalide’s tubulin bound conformation matches well with both the solid-
state and major solution conformations. As shown in Figure 1.11, a dihydropyran ring flip accounts for the only main difference between the two conformations. Given the activity demonstrated by 1.10 and the success of molecular modeling in predicting laulimalide’s conformational preferences, additional development in this area should determine whether it is possible to further simplify the structure and retain activity.

1.5.3 Exiguolide

![Exiguolide Structure](image)

Figure 1.12 – The natural product (–)-exiguolide 1.11.

In 2008, the Cossy laboratory recognized the structural similarity between the bryostatins and the natural product exiguolide 1.11 (Figure 1.12).\(^{64}\) Interestingly, both 20-membered marine polyketides contain analogous structural motifs such as a bis-tetrahydropyran unit and methoxycarbonyl methyldiene group. Ohta and co-workers isolated exiguolide in 2006 from the sponge *Geodia exigua* off the coast of Japan and determined its molecular structure via extensive NMR analysis.\(^{65}\) The complex macrolide inhibits the fertilization of sea urchin (*Hemicentrotus pulcherrimus*) gametes while also
displaying potent *in vitro* anti-cancer activity against multiple human cancer cell lines.\textsuperscript{66} A recent COMPARE analysis suggests exiguolide’s antiproliferative activity may arise through a distinct mechanism of action compared to other anti-cancer agents such as paclitaxel, vincristine, and doxorubicin.\textsuperscript{67} Exiguolide’s unknown mechanism of action along with its structural similarity with the bryostatins has prompted the interest of many synthetic chemists. However, the molecule’s extreme scarcity has limited SAR studies, which in turn has hindered further investigation of this promising anti-cancer agent.

\textbf{Figure 1.13 – Exiguolide analogues developed by the Fuwa laboratory.}
Recently, the Fuwa laboratory complemented existing SAR studies\textsuperscript{68} by focusing on the relationship between exiguolide’s conformation and the molecule’s cytotoxicity.\textsuperscript{69} Since previous studies established the importance of the triene side chain and C5-Z-enoate, Fuwa and co-workers investigated how functional groups on exiguolide’s macrocyclic backbone impacted cytotoxicity. In particular, the Fuwa group focused on removing the C15 and C18 allylic methyl groups since these motifs complicated their overall synthetic strategy. Their synthesis employed a stereoselective domino cross-metathesis/intramolecular oxa-conjugate addition for building the methylene bis-tetrahydropyran core.\textsuperscript{70} Using this advanced intermediate as a divergent starting point, the Fuwa laboratory constructed the 15-desmethyl analogue 1.12 and 15,18-bis-desmethyl exiguolide 1.13 using a synthetic strategy similar to the one utilized in their exiguolide total synthesis. They also prepared the corresponding (16,17-Z)-exiguolide 1.14, as they anticipated this inverted olefin geometry would significantly alter the macrocyclic conformation and serve as an excellent control for gauging the relationship between exiguolide’s conformation and its biological activity.

To evaluate the conformational effects of their structural modifications, the Fuwa laboratory performed NMR-based conformational analyses on 1.11 and each of their newly synthesized analogues. Using 1H-NMR and 2D-NOESY experiments in conjunction with energy-minimized structures generated through MMFF94s calculations, the group concluded that removing the C15 and C18 methyl groups minimally affects the macrocyclic backbone, with 1.11, 1.12 and 1.13 adopting almost identical conformations. Additionally, 1.14 possesses the expected altered conformation, affecting both the C14-C29 domain and bis-tetrahydropyran core (Figure 1.14).
Biological testing of 1.14 showed no growth inhibition against A549 and NCI-H460 cancer cell lines, establishing the importance of maintaining exiguolide’s natural conformational profile. While 1.12 demonstrated comparable activity to the parent natural product (A549 IC$_{50}$ = 3.14 μM vs. 1.66 μM for 11) 1.13 exhibited no activity at >100 μM, leading them to suggest that the C18 methyl group plays some role in activity unrelated to conformation. However, as the removal of either the C15 or C18 methyl should significantly alter local conformational preferences through the loss of A$^{1,3}$-strain, it is possible that the overall conformational preferences of 1.13 were changed to a greater degree than their preliminary analysis anticipates. More detailed conformational analysis employing some of the techniques outlined in this review and the additional synthesis of an 18-desmethy lexiguolide analogue should hopefully clarify the effects of the C15 and C18 methyl groups’ effects on biological activity.
1.5.4 Discodermolide

In 1990, Gunasekera and Longley isolated discodermolide \( \text{1.15} \), a linear polyketide natural product, from the marine sponge *Discodermia dissoluta* (Figure 1.15).\(^{71,72}\) While initially reported as an immunosuppressive agent, follow up studies later established discodermolide as a potent microtubule-stabilizing agent that retains low nanomolar cytotoxicity against paclitaxel-resistant cell lines.\(^{73}\) As expected, this antitumor activity spurred numerous synthetic efforts over the years, with a sizeable number of total syntheses reported to date.\(^{74}\) One noteworthy undertaking by Novartis utilized a mixture of the Paterson and Smith routes, eventually producing more than 60 grams of material for use in clinical studies. However, phase I trials were ultimately halted as a result of severe lung toxicities.\(^{75}\)

![Discodermolide 1.15](image)

Figure 1.15 – Discodermolide \( \text{1.15} \).

Since certain discodermolide structural features may serve a non-essential role for human chemotherapy, removing these extraneous features would create a more accessible
synthetic target and aid the development of more therapeutically benign discodermolide analogues. Several independent laboratories have focused primarily on determining discodermolide’s bioactive conformation for assisting with this analogue design. In 2001, the Smith\textsuperscript{76} and Snyder\textsuperscript{77} groups independently conducted extensive conformational studies with 1.15 using molecular modeling and high-field NMR experiments. Interestingly, these studies showed that discodermolide’s acyclic structure possesses a surprisingly limited conformational profile, with the most favored conformation exhibiting a “hairpin” motif as a result of several key features.

First, the Z-olefins within the polyketide backbone introduce A\textsuperscript{1,3}-strain, stabilizing the C7-C10, C12-C15, and C20-C23 regions of the molecule. The multiple polypropionate substitutions also force discodermolide’s backbone into conformations that minimize syn-pentane interactions, creating two major turns in the C10-C12 and C16-C20 regions. Finally, the six-membered lactone ring interconverts between chair, half-chair, and skew-boat conformations, with a hydrogen-bonding interaction between the C7-hydroxyl and the lactone ring oxygen having some moderate influence over the C5-C7 torsional angles (Figure 1.16). Later transfer-NOE spectroscopy performed in the presence of both unassembled tubulin and assembled microtubules determined that the bound conformation of discodermolide closely matches that of the solution structure, save for some subtle differences in the orientation and shape of the ring.\textsuperscript{78}
The formation of this stereochemically-rich lactone ring requires numerous synthetic steps, reducing the practical efficiency of many discodermolide syntheses. Investigations into the necessity of this structurally complex functionality showed that inversion of the C4 and C5 stereocenters did not lead to a loss in potency.\(^7^9\) Additionally, structure-activity relationship studies revealed that the 2,3-anhydro derivative also retains
nanomolar cytotoxicity. As a result, the Smith laboratory hypothesized that the substituents on the ring orient the lactone in a conformation that maximizes binding interactions rather than serve any binding purpose of their own. Thus, they proposed a more conformationally rigid five-membered scaffold as a simplified substitute for the complex six-membered lactone ring.

Figure 1.17 – Conformationally-simplified analogues of discodermolide. The remainder of the structure is identical to 1.15 has been omitted for clarity.

The synthesis of this simplified analogue required only three steps from one of their advanced intermediates and used cheap and commercially available levulinic acid, giving furan 1.16 in 45% overall yield. The resulting unsubstituted butyrolactone displayed a near 10-fold improvement over 1.15 against the MCF-7 cancer cell line. The subsequent discovery that the C7-hydroxyl played a non-essential role in promoting cytotoxicity suggested that the entire C1-C7 region could be simplified without penalty. Smith and co-workers tested this hypothesis by synthesizing an analogue that replaced the C1-C7 region with a coumarin moiety. Excitingly, the resulting analogue 1.17 also
displayed low nanomolar cytotoxicity against multiple cell lines, despite possessing only 8 of discodermolide’s 13 stereocenters (Figure 1.17). Given this in vitro success, in vivo testing and further development of simplified discodermolide analogues could still yield a promising drug candidate despite the earlier clinical failure of the natural product.

1.5.5 Dictyostatin

![Image of dictyostatin and discodermolide structures]

Figure 1.18 – Comparison of dictyostatin 1.18 and discodermolide 1.15 structures.

In 1994, Pettit and co-workers isolated dictyostatin 1.18, a polyketide structurally resembling a cyclic version of discodermolide, from a marine sponge collected off the Republic of Maldives (Figure 1.18). Like discodermolide, dictyostatin exhibits potent microtubule-stabilizing activity, retains cytotoxicity against multidrug-resistant cell lines, and binds the taxoid site. Given the failure of discodermolide in clinical trials, attention has turned towards dictyostatin as an alternative candidate. To date there have been eight published total syntheses of 1.18, with the Paterson and Curran groups additionally providing a wealth of analogues for testing. Although most of these analogues have exhibited diminished activity, their cumulative biological results have provided a comprehensive picture of dictyostatin’s structure-activity relationships.
Recent pharmacological studies in mice have shown that dictyostatin crosses the blood-brain barrier to effect prolonged microtubule stabilization in the brain, hinting at its potential to treat neurodegenerative disorders such as Alzheimer’s.\textsuperscript{109}

![Figure 1.19 – Overlay of dictyostatin s-trans solution structure (blue) with TR-NOESY conformation (green) and solid-state structure of discodermolide (yellow).](image)

While a crystal structure of dictyostatin has yet to be reported, extensive NMR experiments have provided ample data for conformational studies. The low isolation yield of Pettit and co-workers initially led to an incomplete stereochemical assignment, leaving the structure of dictyostatin unclear until its re-isolation in 2003\textsuperscript{84} and subsequent work by Paterson and co-workers in 2004.\textsuperscript{110} By utilizing both homonuclear ($^{3}J_{\text{H,H}}$) and heteronuclear ($^{2}J_{\text{C,H}}$) coupling constants in combination with NOESY experiments they were able to determine the complete relative stereochemistry, which was later confirmed as the absolute configuration through the concurrent synthetic efforts of the Paterson\textsuperscript{85} and Curran\textsuperscript{86} groups. The Paterson group further combined this NMR data with Monte
Carlo conformational searches to propose a pair of interconverting atropisomers, wherein the lactone adopts either a C1-C2 $s$-trans or $s$-cis orientation. Of the two, the lower-energy $s$-trans structure bears a strong resemblance to the solid state conformation of discodermolide. Later work by Canales et al led to a proposed bioactive conformation of dictyostatin based on TR-NOESY data, which overlays well with the $s$-trans solution structure save for a major torsional change around the C8-C9 linkage (Figure 1.19).\textsuperscript{111}

Given the structural similarities between discodermolide and dictyostatin, there has been substantial interest in developing hybrid analogues to explore potentially shared pharmacophoric elements. Early work by Curran and co-workers pre-dated the elucidation of the true structure of dictyostatin, but still established that simplified macrocyclic analogues could retain modest biological activity.\textsuperscript{102} Following the stereochemical reassignment of dictyostatin the Paterson group revisited the issue by designing and synthesizing their own hybrid molecule 1.19, which computationally matched well with the solid-state conformation of discodermolide.\textsuperscript{94} Initial biological testing revealed that this compound was approximately 10-fold less active than discodermolide, and follow-up studies revealed a further loss in potency against Taxol-resistant cell lines.\textsuperscript{99} After examination of the NMR-derived bioactive conformations of discodermolide and dictyostatin, Paterson and co-workers further refined their hybrid model to include the C1-C5 dienoate region proposed to be necessary for dictyostatin’s binding affinity.\textsuperscript{97} The improved hybrid 1.20 demonstrated nanomolar inhibition against several human cancer cell lines, including Taxol-resistant ones. Interestingly, thermodynamic experiments have found that the acyclic discodermolide’s binding to
tubulin is entropically more favorable than the cyclic dictyostatin, suggesting that the latter’s activity may be unrelated to any relative lack of conformational freedom.\textsuperscript{112}

Figure 1.20 – Dictyostatin-discodermolide hybrid molecules.

Recently, Snyder and co-workers disputed the conformational picture of dictyostatin as incomplete, arguing that its structure is not rigid enough to support only two conformational families.\textsuperscript{113} Using NAMFIS to re-evaluate the methanol-$d_4$ NMR data set, they identified sixteen conformations ranging in population from 11.4\% to 2.5\%; although the previously proposed major solution conformation was not substantially populated, they did find a roughly equal distribution of C1-C2 \textit{s-trans} and \textit{s-cis} orientations amongst the set. They also acquired NMR data in DMSO-$d_6$ in order to investigate solvent effects with more quantitative data, ultimately finding fifteen conformations with the \textit{s-cis} being primarily represented. Among the conformations from the two solvents they found that one of their DMSO structures was a good fit for the TR-NOESY data reported by Canales \textit{et al.} Subsequent docking of this compound to tubulin uncovered a competitive binding pose that is radically different from the previously reported one, yet compatible with several aspects of known dictyostatin SAR.
While this latest work significantly expands on the conformational profile of dictyostatin, energy calculations on the newly-obtained conformations would have been beneficial in gauging their feasibility. The authors are correct in asserting that force field energies aren’t necessarily accurate but neglected to explore higher \textit{ab initio} methods to assess relative strain energy, particularly given the prevalence of the higher-energy s-cis form amongst their conformations. Additionally, their binding pose suffers from both the low resolution of the crystal structure used and the inherent inaccuracies of rigid docking to a binding pocket that originally contained a different ligand. Given the significant discrepancies between reported binding poses, additional experimental or crystallographic work is necessary to resolve the issue of dictyostatin’s bound conformation.

1.5.6 Trienomycin A

The ansamycin natural products are a broad class of polyketides that exhibit potent antibiotic and anti-neoplastic activity, including the well-known rifamycins. In 1985, Umezawa and co-workers isolated the structurally related natural products, trienomycins A-F, from a culture broth of \textit{Streptomyces} sp. No. 83-16.\textsuperscript{114} Subsequent biological testing revealed that the trienomycin class displays potent in vitro cytotoxicity against multiple cancer cell lines.\textsuperscript{115} Of the six known ansamycins, trienomycin A \textbf{1.21} demonstrated the strongest cytotoxicity with an IC\textsubscript{50} value of 0.01 μg/mL against L-5178Y murine leukemia and human PLC hepatoma cell lines. The trienomycins primarily consist of a 21-membered macrolactam ring adorned with a non-redox active phenol, an (\textit{E},\textit{E},\textit{E})-triene, four stereocenters, and an \textit{N}-acylated D-alanine side chain, varying only in the composition of the terminal side chain functionality (Figure 1.21).
Unfortunately, the clinical potential of the trienomycins represents an underexplored avenue as a result of supply issues and minimal structure-activity relationship studies. Funayama and co-workers prepared four semi-synthetic trienomycin A analogues and concluded that the C13-OH, the triene subunit, and the N-acylated side chain were essential for potent cytotoxicity. However, methylation of the free phenol produced an analogue equally active to the parent natural product. These classical SAR studies reveal the difficulty in correlating molecular functionality to biological activity within a polyketide macrocyclic framework.

The Blagg laboratory addressed these difficult issues by employing a rational, conformational approach for designing simplified analogues of trienomycin A. After generating low-energy conformations of the semi-synthetic analogues with SYBYL, Blagg and co-workers showed that while Funayama’s inactive derivatives exhibited significant conformational differences from trienomycin A, the active methyl ether

![Figure 1.21 – Trienomycins A-F.](image-url)
analogue retained the parent geometry. The Blagg group hypothesized that mimicking the conformational preferences of 1.21 properly orients the critical phenol unit and N-acylated side chain within trienomycin’s binding pocket, thereby promoting tight binding. Interestingly, modeling suggested that removing the C13-OH, C12-Me, and C3-OME functional groups minimally perturbed the macrolide’s overall conformational preferences. Moreover, additional removal of the C4/C6 alkenes and inversion of the C14-(Z)-alkene geometry produced monoene A 1.22, which SYBYL calculated would adopt a conformation strongly resembling that of 21 (Figure 1.22).

Figure 1.22 – Overlay of trienomycin 1.21 (green) and monoene A 1.22 (yellow) SYBYL-generated structures.

The synthesis of 1.22 required only 13 steps from readily available starting materials. This concise route allowed the multi-gram preparation of not only monoene A but also monoene E 1.23, which SYBYL calculations indicated would act as a negative control due to its low similarity score. As their model predicted, the conformationally similar 1.22 displayed potent anti-proliferative activity (0.47 nM, MCF-7 cell line) while
the lowest similarity structure 1.23 displayed no significant activity. Buoyed by these results, the Blagg laboratory embarked on an SAR investigation of the side chain using the new simplified macrocycle as a framework, which effectively found that alterations to the α-methyl and alkyl amide functionalities were not well-tolerated.

![Figure 1.23 – Monoene A 1.22 and E 1.23 analogues.](image)

Blagg’s approach shows that a simplified analogue of the synthetically-complex trienomycin A could be used for faster determination of SAR and could potentially be used for elucidation of the biological target. However, given the near-certainty that 1.22 is significantly more flexible than the parent natural product, it is possible that trienomycin’s mode of cytotoxicity may be more dependent on recognition of the side chain rather than the macrocycle.

1.5.7 Spongistatin

The spongistatin class of natural products was first reported in the early 1990s by three separate groups.\(^{119-121}\) Isolated from an Eastern Indian Ocean sponge of the genus *Spongia*, members of this family possess potent antineoplastic properties, with spongistatin 1 1.24 in particular displaying an average IC\(_{50}\) value of 0.12 nM against the
NCI panel of 60 human cancer cell lines. The activity observed during cell growth inhibition assays led Hamel to propose that the spongistatins’ mode of action involves tubulin binding, and hypothesized a “polyether” binding site on β-tubulin near the vinca domain to account for its competitive inhibition of maytansine and rhizoxin.

\[\text{Figure 1.24 – The spongistatin class of natural products.}\]

The complex structure and potent activity of the spongistatins attracted considerable attention from laboratories interested in exploring its potential as a chemotherapeutic agent. Given their scarcity and difficulty to obtain from nature,
significant resources have been devoted towards their creation in the laboratory, with several total syntheses of $1.24^{125-131}$ and $1.25^{132-136}$ reported to date. Among these, the Smith laboratory took possibly the most ambitious approach in their fourth-generation synthesis of spongistatin 1. Utilizing a multi-component dithiane union tactic, Smith and co-workers assembled multiple, key fragments in a scalable fashion to yield a remarkable 1.009 g of material.\textsuperscript{137}

Having obtained enough material for preclinical development, the Smith group turned towards the task of identifying the moieties necessary for cytotoxicity through SAR studies.\textsuperscript{138} Given the difficulty in constructing analogues even through their optimized route, the Smith group investigated whether the entire structure of the flexible spongistatin skeleton was necessary for biological activity. Based on the structural features common amongst the natural spongistatins and results from several early analogues,\textsuperscript{139-142} the Smith group hypothesized that the spongistatins’ western perimeter represented the main pharmacophoric motif. Furthermore, the Smith laboratory believed that as long as the C37-C39 dihedral angles between the E and F rings maintained a proper orientation, replacing the remainder of the complex molecule with a simplified linker would still produce a molecule with potent activity.

To facilitate the design of such a linker, the Smith group used molecular modeling as a means of determining spongistatin’s conformational preferences.\textsuperscript{143} Despite a previously reported solution conformation,\textsuperscript{144} their calculations and resulting analysis of the macrocyclic torsional angles using an expansion of the Taylor laboratory’s polar map methodology\textsuperscript{145} indicated a fairly rigid western perimeter. The remainder of the molecule exhibited high flexibility and adopted multiple conformations. To combat the problems
they perceived with NMR-constrained molecular dynamics (high-energy conformations) and molecular mechanics (inaccurate energy ordering), Smith and co-workers developed a hybrid software method called Distribution of Solution Conformers (DISCON).\textsuperscript{146} DISCON uses NMR-derived interproton distances and torsional angles for determining the most populated families from a computationally generated library of conformers, utilizing hierarchical clustering and a genetic algorithm to avoid the overfitting problems that can arise from NMR-based structural optimization in methods like NAMFIS. A more detailed breakdown of Smith’s analysis and their DISCON software will be covered in Section 2.5.

![Figure 1.25 — Retained ABEF conformational structure of spongistatin 1 (green).](image)

From this DISCON analysis, the Smith laboratory found that \textbf{1.24} existed as an ensemble of four major solution conformations, which when overlaid showed that the western perimeter of the macrocycle maintained a common conformation reflecting an
internal hydrogen bonding network in support of their earlier modeling work (Figure 1.25). To test their linker hypothesis, the Smith group designed an analogue that would possess only the E and F ring systems, ultimately deciding on a biaryl ether tether that would maintain the conformational twist observed during conformational analysis. Synthesis of this molecule (1.26) ultimately revealed that the strain energy imparted during macrocyclization opened the E ring system during a global deprotection.

![Figure 1.26](image1.png)

Figure 1.26 – Simplified spongistatin analogues 1.26 and 1.27 developed by the Smith lab.

To combat this issue, the Smith group redesigned their analogue to instead contain a more flexible polymethylene tether. Since previous attempts at replacing the eastern perimeter with a simple tether failed, they also incorporated an internal hydrogen bond acceptor within the molecule as a means of lowering the overall flexibility. Following synthesis of this second-generation analogue 1.27, DISCON studies found that its major conformations retained the rigidity earlier observed in the ABEF ring system.
and even imparted a certain degree of restraint on the side chain. Incredibly, biological testing found that 1.27 displayed nanomolar cytotoxicity against several cancer cell lines and a subsequent cell cycle analysis concluded that the molecule also shared the same microtubule-destabilizing activity as spongistatin 1.\textsuperscript{147} This work demonstrates how the pharmacophoric elements of a complex natural product are likely to reside on a conformationally-rigid area, and that replacing flexible regions with simplified tethers is a viable design strategy.

1.6 Conformationally-Restricted Analogues of Polyketide Natural Products

Although flexible polyketides adopt preferred conformational profiles, typically only one of these structures will resemble the ideal bound conformation. As each member of a conformational ensemble has an opportunity for protein binding equivalent to its population in solution, altering their relative populations in solution is likely to affect biological activity. Therefore, analogues that alter a molecule’s conformational preferences to favor either a single conformer or family of conformers could be used as a method for identifying the bioactive conformation in the absence of bound crystallographic or spectroscopic data. Additionally, modifications that preferentially favor the bound structure may lead to various improvements over the unrestricted natural product.
1.6.1 Epothilone

Figure 1.27 – The epothilone class of natural products.

In 1987, Reichenbach and Höfle isolated the epothilone class of natural products from the soil-dwelling myxobacterium *Sorangium cellulosum*.\(^{148}\) Nearly a decade later, a group at Merck Research Laboratories discovered that the family not only possessed microtubule-stabilizing properties, but also exhibited similar biological properties as the well-known drug, paclitaxel.\(^{149}\) The epothilones also possessed improved water solubility and retained activity against P-glycoprotein-expressing multiple drug resistant cell lines, including those resistant to paclitaxel.\(^{150,151}\) The intriguing biological activity, pharmacological properties, and accessibility through fermentation spurred significant interest towards developing the epothilones as an anticancer agent, with several semi-synthetic epothilone-type compounds entering clinical trials and one, an epothilone B lactam analogue, being approved for use in humans as the drug Ixempra (ixabepilone) in 2007.\(^{152}\) In addition to clinical development the epothilones received a staggering amount
of attention from the synthetic community, which has been thoroughly covered in reviews elsewhere.\textsuperscript{153}

Figure 1.28 – Conformational analysis of the C1-C8 and C11-C15 regions of 1.29.

Early displacement experiments revealed that the epothilones compete for the same tubulin-binding site as paclitaxel, which prompted interest in their three-dimensional structure and potential homology between the two agents. However, epothilone’s unknown relative stereochemistry plagued early efforts to elucidate the molecule’s solution conformational preferences.\textsuperscript{154} In 1996, Höfle and co-workers used X-ray crystallography as a way of assigning the relative and absolute stereochemistry of epothilone A 1.28.\textsuperscript{155} Using a combination of vicinal coupling constants and observed NOEs, they concluded that the solution conformation of 1.28 was highly similar
compared to its solid-state conformation. However, more detailed NMR studies in multiple solvents showed that several conformational families existed in solution. Solution studies conducted by the Taylor laboratory showed that the epothilone macrolide core possessed two separate solution conformations in the C1-C8 region, with the major conformation resembling the previously reported solid-state structure. A similar analysis of the C11-C15 region demonstrated that flexibility in this region provided additional conformational families, which further complicated the overall picture in solution (Figure 1.28). A more detailed discussion of the methodology used in these solution studies is presented in Section 2.4.1.

Figure 1.29 – Conformational analysis of the C11-C15 region within Taylor’s C14-substituted analogues 1.30 and 1.31.

In an effort to distinguish between several conformational families in solution, the Taylor laboratory investigated multiple structural analogues of the epothilones. They
hypothesized that the incorporation of additional A$^{1,3}$-strain or syn-pentane interactions would alter the relative population of solution conformers, anticipating that the biological activity of the resulting analogues would indicate the bound conformation. To test this theory, they synthesized two diastereoisomeric C14-methyl substituted epothilone analogues, (S)-14-methyl 1.30 and (R)-14-methyl 1.31 (Figure 1.29). They postulated that the introduction of a methyl substituent would alter the relative population of the two observed C11-C15 solution conformational families via A$^{1,3}$-strain. Interestingly, biological testing of both compounds revealed that 1.30, which resembled the solid-state conformation, retained activity while its C14 epimer 1.31 lost all cytotoxicity. In a subsequent study, installation of a C14-methoxy group in the same stereochemical orientation resulted in an epothilone D analogue with significantly improved activity over the unmodified compound.

The conformational preferences of the C11-C14 region also explained the puzzling results observed by the Danishefsky laboratory in their synthesis of (E)-9,10-dehydro epothilone B 1.32. In the last step of the synthesis, Danishefsky and co-workers found that epoxidation of the (E)-9,10-dehydro epothilone D alkene precursor proceeded with an unusual preference for the α-face, which they rationalized on the change in conformational preferences induced by the 9,10-unsaturation. The Danishefsky group also attributed the molecule’s significant increase in potency to conformation as well, since the unsaturation rigidifies the C9-C10 torsional angle and stabilizes the C25 methyl through minimization of A$^{1,3}$-strain. The resulting conformation resembles the solid-state and major solution conformations (Figure 1.30). In addition, analogues probing the conformational families within the C5-C10 region supported a bound
conformation mimicking the Höfle solid-state conformation.\textsuperscript{159} Intriguingly, transfer-NOE and cross-correlated relaxation NMR studies with monomeric tubulin also supported this conclusion.\textsuperscript{160} Figure 1.30 – (\textit{E})-9,10-dehydroepothilone B \textbf{1.32} and its major solution conformer.

In 2004, Nettles and co-workers published an electron crystallographic structure of epothilone A bound to \(\alpha,\beta\)-tubulin in zinc-stabilized sheets, which reported a ligand conformation that diverged significantly from conformations derived via NMR spectroscopy and the previously mentioned analogue studies.\textsuperscript{161} One potential reason for this may be that the crystal structure had insufficient resolution to define epothilone’s conformation, forcing the investigators to supplement their work with a NAMFIS-derived structure. Given these contradicting results, Erdélyi and co-workers embarked on a study of epothilone’s conformational preferences in aqueous solution as a way of rectifying the dissimilarity between the reported bound and free structures.\textsuperscript{162} They also addressed the
conformational role of the 3-hydroxy functional group, as two analogues lacking this moiety, C3-deoxy 1.33 and 2,3-dehydro 1.34, remained active despite its proposed H-bonding interaction within the tubulin-binding site.\textsuperscript{163,164} Their analysis revealed that the previous NMR-derived conformation of unbound epothilone A was a probable component of their aqueous conformational ensemble, and found that the EC-derived conformation was not significantly populated in water. Moreover, Erdélyi et al observed that the C3-deoxy analogues possessed similar conformational preferences and that removal of the C3-hydroxy both with and without the incorporation of a \textit{trans} double bond between C2 and C3 does not alter the preferred overall macrolide conformation.

![Figure 1.31 – C3-deoxyepothilone 1.33 and 2,3-dehydroepothilone 1.34.](image)

Recently, Prota and co-workers resolved the issue of epothilone’s bioactive conformation by obtaining a high-resolution X-ray co-crystal structure of epothilone A bound to \textit{α,β}-tubulin.\textsuperscript{165} They accomplished this extraordinary feat through the incorporation of stathmin-like protein RB3 and tubulin tyrosine ligase to the complex. This resulted in a protein-ligand complex with high enough resolution to directly define the conformation of bound epothilone A, which matched well with both the bound and solution structures reported in earlier NMR and analogue studies (Figure 1.32). The
complicated history of the epothilones shows how accurate prediction of a polyketide’s conformational preferences can lead to the prediction of its bioactive structure, which is of critical importance when a well-defined protein co-crystal structure cannot be easily obtained.

![Figure 1.32 – Overlay of major solution conformer (green) and tubulin-bound structure (orange) of 1.28.](image)

1.6.2 Geldanamycin

As mentioned previously, the ansamycin natural products have received considerable attention as potential chemotherapeutic agents for the treatment of several diseases. In particular, the benzoquinone ansamycin, geldanamycin 1.35, exhibits potent anti-cancer activity by binding the ATP site on heat shock protein 90, a molecular chaperone involved in the folding and maturation of proteins upregulated in tumors. In 1970, DeBoer and co-workers isolated geldanamycin from *Streptomyces hygroscopicus* var. geldanus, with its molecular structure being elucidated soon thereafter by the
The 21-membered polyketide contains a trisubstituted benzoquinone ring, six stereocenters, and an α,β-unsaturated dienamide. As expected, this structurally daunting macrocyclic framework has challenged numerous synthetic chemists.

Figure 1.33 – Geldanamycin 1.35 and its analogues.
In 2002, Andrus and co-workers accomplished the first total synthesis of \textit{1.35} utilizing glycolate aldol chemistry and a late stage oxidation for construction of the challenging quinone unit.\textsuperscript{168} Six years later, the Panek laboratory also completed its total synthesis in only 20 linear steps and 2\% overall yield.\textsuperscript{169} Biosynthetic production would later provide large quantities of geldanamycin and expedite the synthesis of multiple semi-synthetic analogues, with most modifications occurring at the C17 quinone position. Structural alterations at this position included addition-elimination reactions with several nucleophiles such as alkoxides, hydroxide, phenoxide, and amines. The C17 position was targeted as a means of improving the solubility of geldanamycin; however, even the most promising allyl amine derivative \textit{1.36} still retained significant toxicity. Literature reports suggest the hepatotoxicity of geldanamycin arises via the reaction of biological nucleophiles such as glutathione with the C19 quinone.\textsuperscript{170-172} Curiously, geldanamycin possesses two condition-dependent conformations; an extended trans-amide solution/solid-state conformation supported by X-ray crystallography\textsuperscript{173} and NOE experiments,\textsuperscript{174} and a closed ‘C-clamp’ cis-amide bound conformation arising from protein co-crystallography.\textsuperscript{175-177} This discontinuity sparked a number of different studies, which disagreed on either the energetic barrier of the isomerization or its overall necessity for binding and inhibition.\textsuperscript{178-181} Kitson and co-workers hypothesized that substitution at geldanamycin’s C19 position could potentially kill two birds with one stone, blocking glutathione’s nucleophilic addition while also forcing the unbound macrocycle into a \textit{cis}-amide conformation. In doing so, they believed they could attenuate geldanamycin’s toxicity issues while also investigating the conformational requirements of its binding.\textsuperscript{182}
Starting from commercially available geldanamycin, Kitson *et al* prepared a wide variety of C19-substituted analogues. As modification of this site through nucleophilic addition possessed limited scope, they maximized analogue diversity by employing Stille coupling reactions with 19-iodogeldanamycin 39, which was readily obtained in a single step from 35. After successfully synthesizing 19-methylgeldanamycin 40 in 86% yield, they attached several electron-rich and electron-deficient aromatic groups to the C19 position in good to excellent yields. Since 17-substituted derivatives such as 36 and 37 showed earlier clinical promise, Kitson and co-workers also synthesized multiple analogues of their 19-substituted geldanamycin derivatives.

![Figure 1.34 – Overlay of solid state 1.42 (orange) with protein-bound 1.35 (gray).](image)

Interestingly, the 19-substituted analogues displayed higher polarity than geldanamycin. Kitson and co-workers undertook conformational NMR studies as a
means of examining whether observed polarity resulted from the hypothesized \textit{trans}- to \textit{cis}-amide isomerization. While 2D NMR work on geldanamycin and the 17-amino derivative \textbf{1.38} suggested a preference for the \textit{trans}-amide conformation, the 19-substituted compounds revealed substantial $^1$H and $^{13}$C chemical shift differences that implied a major change in the chemical environment throughout the macrocycle. Quantitative NOE studies performed on 19-phenyl-AAG \textbf{1.41} strongly suggested that the \textit{cis}-amide conformation formed predominantly in solution. The resulting ‘C-clamp’ form seen in protein crystallography studies moves the hydrophobic surface area into the heart of the structure while exposing hydrophilic moieties to the solvent resulting in the observed increase in polarity. X-ray crystallography studies on 19-substituted analogue \textbf{1.42} confirmed these results, with the resulting crystal structure overlaying well with that of protein-bound geldanamycin (Figure 1.34).

Using \textit{N}-acetylcysteine methyl ester as a model for glutathione attack, Kitson and co-workers determined that while geldanamycin and 17-amino derivatives will react under basic conditions, 19-methyl and 19-phenyl analogues do not. To determine the biological implications of blocking nucleophilic attack at C19, they tested several 19-substituted derivatives using normal endothelial and epithelial cells. The resulting data showed that these analogues exhibit less toxicity than their parent compounds (\textbf{1.35-1.37}) in all tested cases. Protein crystallography studies confirmed that the 19-substituted compounds bind in the ATP site on Hsp90, although the added steric bulk alters the quinone ring’s bound position thus penalizing the molecule’s binding affinity. This work reiterates how simple alterations to a polyketide’s structure can have profound implications to its conformational preferences.
1.7 Conclusions and Future Outlook

Polyketide natural products possess a wide range of structural features that play a role in defining their conformational preferences through fundamental steric, electronic and electrostatic interactions. These evolutionarily driven features confer the overall structure with defined conformer populations while retaining backbone flexibility, which likely provides energetic assistance during protein binding and improved physical properties such as solubility and membrane transport. Computer-based molecular modeling can provide an accurate representation of the conformational preferences of even the most complex polyketide structures. The resulting library of potential conformers can be ordered energetically through the use of solution NMR-derived coupling constant and NOE data. As is clear from the case studies covered within this review, protein-bound conformations determined by X-ray crystallographic analysis often show significant similarities to low-energy conformers populated in solution.

From one perspective, conformational mimics can be designed to maintain native conformer populations with simplified structural features. This design strategy can address supply issues for clinically bound polyketides in service of function-oriented synthesis principles. Alternatively, the addition of structural features that attenuate conformational preferences can provide insight into the bound conformation when coupled with biological activity studies. In either case, it is clear that investigators cannot disregard the conformational requirements of a polyketide’s biological activity in their analogue design programs.
CHAPTER 2:

FROM NAMFIS TO DISCON: TOOLS FOR CONFORMATIONAL ANALYSIS

2.1 Purpose

The purpose of this chapter is to outline the tools used for analyzing molecular shape in the course of a conformational analysis. Crystallography, NMR analysis and \textit{in silico} molecular modeling will be covered, followed by a discussion of several hybrid methods that have been used to varying degrees of success in solution conformational studies of polyketides. Finally, a complete outline of the current conformational analysis workflow will be presented, including the various updates to the procedures implemented as part of this work.

2.2 Tools for Analyzing Small Molecule Shape

Determining molecular shape comes down to the accuracy of the model one can build, which makes the technique dependent on tools that reflect both quantitative and qualitative measurements of conformation. For example, while a plastic molecule kit provides a wonderful qualitative means to gauge how steric interactions control conformation, experimental evidence is needed to deliver the quantitative data needed for the full picture. While there are a variety of methods available to probe structure and conformation, few offer enough utility and detailed information to stand on their own.
The most popular techniques used for determining molecular conformation are crystallography, nuclear magnetic resonance (NMR) and molecular modeling.

2.2.1 Crystallography

X-ray crystallography represents the apparent gold standard of the conformation world. Information provided by this technique comes from the growth of a solid crystal of the small molecule or protein in question, in which the crystallized atoms deflect beams of X-rays into multiple directions. By analyzing the resulting diffraction patterns, one can produce a three-dimensional image of the electron density within the crystal itself through back-calculation, which in turn allows the approximate positions of the atoms (typically non-hydrogen) inside that density to be determined. When combined with a computer model, this also allows for the determination of connectivity, bond angles and lengths, ultimately providing a fairly accurate conformation of the structure in question.

However, the solid-state nature of crystallography can be limiting when dealing with flexible molecules such as polyketide natural products. Flexibility in crystallography is generally associated with higher B factors (temperature factors), which can spread electron density and introduce disorder to the unit cell.\textsuperscript{184} Also, even if an ideal crystal is achieved the data derived from it will provide a single, static conformation of a given structure, which offers little utility to studies attempting to correlate biological activity to conformation. Because of the need to grow a macroscopic crystal from a pure sample, this further opens the possibility of the molecule’s conformation being influenced by the crystal packing arrangement, which can be very different from a solution environment. As a result, single-molecule crystallography of a flexible biologically-active polyketide
leaves unanswered the question of whether or not the solid-state conformation is related to the bioactive conformation. Addressing this issue requires the co-crystallization of the molecule in question with its protein target, which raises additional issues; not all proteins can be crystallized reliably or quickly, and a low-resolution structure (as in the example of the initial epothilone-tubulin complex, which is detailed in Section 5.2) opens the door to conflicting solutions.

2.2.2 NMR

Nuclear magnetic resonance (NMR) provides a flexible alternative to crystallography, both metaphorically and literally. NMR also provides data on molecular conformation with the advantage of a far lower barrier of entry. One can obtain data relatively quickly in a variety of different solvents instead of a single, slow-growing solid-state environment. NMR data also has the advantage of being dynamic, allowing for analysis of a molecule’s flexibility instead of the static picture provided by crystallography.

NMR data for conformation is generally obtained in the form of $J_{1,3}$ coupling values and nuclear Overhauser effect (NOE) cross-peaks, with the former providing information on bond dihedral angles and the latter offering interatomic distances. Though this data cannot offer atomic coordinates like X-ray crystallography, NMR often provides the information necessary to determine multiple conformations and the relative populations of each. However, deconvolution of this data into individual structures provides for a challenge. The time-averaged nature of NMR means that all possible conformations are represented in a single spectrum, making it difficult to discern the individual contributions from multiple structures. Additionally, data from coupling values
is not necessarily definitive; the magnetic splitting is the result of the absolute angular orientation of two protons, which means that a number of different rotational values can result in the same observed coupling constant (Figure 2.1). A combination of model kits and intuition can be used to narrow the scope of possibilities, but this process can be extremely time-consuming and could miss potential conformational families.

![Diagram of molecular structures with dihedral angles and coupling constants](image)

**Figure 2.1** – Approximate $^3J_{H-H}$ coupling constants for a given dihedral angle between two protons, depending on heteroatom substitution of $R_{1-4}$

### 2.2.3 Molecular Modeling

A third approach to determining conformational preferences eschews experimental data in favor of predictions based on molecular structure. Computational molecular modeling uses *in silico* methods to reduce a molecule’s structure to numbers, with changes in conformation correlating to calculated changes in potential energy through strain, steric and electronic effects. By subjecting a given structure to a search of potential conformations, one can obtain a library of theoretical conformers that will reflect the available conformational space within a given energy window. A variety of different programs and methods exist to perform these calculations at varying levels of accuracy and expense. One of the key benefits of molecular modeling is that it can be
performed without a sample of the molecule in question, allowing for a reasonable level of theoretical insight to preclude and guide a potentially costly experimental process.

Without the data provided through experiment like crystallography and NMR, molecular modeling is highly reliant upon the quality of the mathematical parameters used in a given method. For most methods involving small molecules, this comes down to the force field model used for calculating the potential energy of a given conformation. Molecular mechanics force fields approximate the quantum mechanical energy surface with a classical mechanical model, which decreases the computational cost of simulations by several orders of magnitude. These fields are generally subdivided into classes, with common ones for polyketide conformational analysis either in the class I (AMBER, OPLS, CHARMM) or class II (MM2/3, MMFF94) variety; class III fields, which include polarizability, have not been significantly explored in polyketides. Class I fields represent the sum of bonded (bond stretching, angle bending, torsional terms, improper dihedrals) and non-bonded (Coulomb interactions) energy terms. Class II fields also include additional cross-terms that reflect the coupling between adjacent bonds, angles, and dihedrals, as well as cubic potential energy terms that allow for a more accurate reproduction of QM potential energy surfaces. These additional parameters increase the accuracy of the calculation at the cost of increased computational time, which makes class II fields well-suited for evaluating the conformation of small molecules (such as polyketides) but unwieldy for larger structures such as proteins.

Additionally, the choice of method can have a drastic effect on the makeup of the results. Both Monte Carlo molecular mechanics searches and molecular dynamics simulations will yield a considerable number of local minima along the potential energy
surface. However, Monte Carlo searches offer several advantages over molecular
dynamics when it comes to conformational searches in polyketides, particularly with
respect to efficiency. If the molecule in question possesses any significant conformational
energy barriers then MD simulations, which generate conformations through simulation
of the physical motion of the molecule over time, can become trapped between several
low energy conformations and will not provide a full conformational sampling.
Conversely, Monte Carlo searches are randomized, altering individual torsion angles
before reconnecting and energy-minimizing the structure, and therefore can easily cross
energetic barriers. Additionally, it does not require computationally-costly energy
derivatives, explicit solvents, thermostats, or baryostats, which means that even low-
powered machines should be able to comprehensively cover the conformational space of
a complicated polyketide.

While molecular mechanics methods are generally used for conformational
searches due to their speed, their relative inaccuracy means that other computational
methods are needed for accurate conformational energy predictions. This is usually
accomplished through Density Functional Theory (DFT) calculations, which uses
quantum mechanical methods to determine a far more accurate potential energy value for
a given conformation. However, these calculations are extremely time-consuming, and
are better suited to refining and optimizing individual structures instead of exploring
overall conformational potential. Overall, in silico methods offer valuable perspectives
for conformational analysis, though somewhat limited to theoretical refinements to
experimental data.
2.3 NAMFIS Analysis of Molecular Models

Each of the methods detailed above possesses significant drawbacks when used on their own. X-ray crystallography yields only a single static conformation, NMR provides averaged data that can inaccurately depict a single virtual conformation rather than interconverting conformers if poorly analyzed, and unconstrained molecular mechanics offers only theoretical data. As such, a quantitative analysis of molecular shape typically requires using at least two of these techniques in tandem to provide a detailed picture of a molecule’s conformational preferences. Since X-ray crystallography is better suited for solid state analysis, combining NMR with molecular modeling presents the best strategy for detailing the conformational preferences of a flexible molecule in solution.

One approach in unifying these methods is a method that uses spectroscopic data to perform an iterative refinement of molecular modeling data to yield representative conformers. This sort of analysis was developed by Cicero and co-workers in 1995, termed NMR Analysis of Molecular Flexibility in Solution (NAMFIS). Originally performed on peptide substrates, they began by performing a conformational search using MEDUSA (Multiconformational Evaluation of Distance information Using Stochastically constrained minimization Algorithm). The resulting conformers were clustered statistically to yield a final library. Following this, they used their NAMFIS algorithm to evaluate their refined conformers with respect to observed NMR data, specifically NOE and $J_{1,3}$ coupling constants, from which they determined representative conformations and the relative molar percentage of each.
2.3.1 Laulimalide

As mentioned in Section 1.5.2, the Snyder lab employed NAMFIS methodology in their study on laulimalide’s conformational preferences,\textsuperscript{59} which serves as a useful primer on the application and drawbacks of this approach with regard to polyketide natural products. Their analysis began with a molecular mechanics-based Monte Carlo conformational search along with the implicit GB/SA water solvation model and the AMBER, MMFF and MM3 force fields. Using the available unbound crystal structure as their starting geometry,\textsuperscript{57} they performed a 75,000-iteration search with a 26 kJ/mol energy cutoff, which they deemed necessary in order for the search to re-find the starting geometry an appropriate number of times (>12). From this search, they obtained a library of over 15,000 structures, which were reduced to approximately 5,000 after XCluster\textsuperscript{195} refinement to remove duplicates.

Their experimental refinement utilized eight $^3$J$_{H-H}$ coupling constants and 79 NOE crosspeaks, which they obtained from $^1$H NMR with DMSO as a solvent. Use of these parameters for NAMFIS refinement cut their library of 5,000 down to 22 unique conformers, with estimated populations ranging from 1.3% to 17.2%. Each of these conformers were then evaluated for single-point DFT energies using B3LYP/6-31G*, and structures containing energies greater than 20.92 kJ/mol (5 kcal/mol) above the global minimum were further eliminated to yield a final library of 15 conformers. These were grouped into five conformational families based on general trends in macrolide shape and side chain placement, termed supine (33.7%), convex (30%), stretch (20.8%), concave (8.2%) and cobra (7.2%), respectively. The crystal structure of laulimalide fell into the most populated of these, supine.
While this analysis provides a thorough look at the conformational space that laulimalide can occupy, their methodology raises some issues with their conclusions. The first issue comes from the final winnowing of their NAMFIS results via DFT. These *ab initio* calculations found that several structures possessed energies far in excess of the global minimum (6.4-12.8 kcal/mol, or 26.8-54.4 kJ/mol), yet the NAMFIS refinement of their conformational library predicted these combined structures to account for 19% of the total solution population (seven structures, ranging from 4.8% to 1.3%). As solution population decreases precipitously as conformations increase in potential energy, it is extremely unlikely that any of these structures would have been observable in the experimental data. These DFT energies may have been anomalous due to their methodology, as they took single point energies of MMFF-minimized structures instead of optimizing the geometry in DFT, but they merely dismissed the structures from their pool without considering their potential contributions to the conformational picture.

The second major issue with their analysis involves their treatment of potential *syn*-pentane and *A*¹,²-strain interactions, which were present in several of their predicted conformers. They argued that contacts such as these do not necessarily disqualify a conformation from being a viable solution structure as long as the energy penalty is offset elsewhere. To test this, they manually manipulated the offending torsional angles to reflect both folded and extended angles and comparing the resulting conformers via DFT calculations. Despite incurring significant energy penalties in several structures, they concluded that the additional intramolecular hydrogen bonding offered as part of these conformations was enough to offset it. While they are correct in asserting that a penalty of 1-2 kcal/mol is insufficient to eliminate a conformation, their methodology here should
have mentioned the NOE crosspeaks that supported their conclusion instead of relying solely upon theoretical means. Additionally, the intramolecular hydrogen bonding contacts they cite for energy offsets may not be present in a solution environment, depending on the solubility needs of laulimalide.

Figure 2.2 – Superposition of the six conformers from the “supine” family. Figure taken from Thepchatri et al, *J. Am. Chem. Soc.*, **2005**, *127*, 12838.

Finally, their assignment of structures to motifs is not particularly coherent. For example, the most-populated “supine” family contains only six members, yet the overlap of these structures yields a confusing jumble of shapes (Figure 2.2). Additionally, members within a family can possess drastically different structural features, which is particularly evident in the “convex” family. Despite possessing only three members, conformers 1, 7, and 18, the three have inconsistencies in the number of predicted hydrogen bonds (two in 1 and 7, none in 18) and *syn*-pentane interactions (present in 1 and 18, absent in 7).
Snyder’s analysis concludes that laulimalide is a highly flexible molecule, occupying a fairly wide range of conformations in solution. However, despite the impressive amount of conformational data generated by their approach, the incongruities present in their motifs and the necessary elimination of high energy structures found by NAMFIS provides an ambiguous picture of the molecule’s preferred conformations. This in turn would make it difficult to design analogues to elucidate the bioactive conformation, as it lacks a comprehensive analysis of local acyclic conformational preferences.

2.4 Polar Map Analysis of Molecular Models

One alternative to using an integrated method for determining solution conformations is to perform a full molecular modeling analysis prior to NMR refinement. In doing so, one can determine the likely major conformational families without needing the conformer elimination and recalculation in NAMFIS. However, this approach means dealing with a potentially massive amount of data from conformational searches, specifically the wide array of unique conformers that can be generated within a relatively narrow energy window. Therefore, using this methodology requires a way to evaluate a large number of structures, with the goal of ultimately sorting them into representative families based on similar conformational preferences.

Our approach to this problem involves breaking down a conformation into a collection of its individual backbone bond rotations, which we represent using individual dihedral angles. By combining this data with the relative potential energy of a conformation (as determined by the force field used in the conformational search) and plotting the entire library onto a polar graph, we can obtain a set of polar maps that reflect
the conformational preferences of the molecule. By analyzing these plots, we can define conformational families through observation of rigid and flexible regions within a molecule while also providing a reasonable measure of conformational preferences with regard to potential energy, all without requiring any elimination of conformers from the search. Once we have an analyzed set of conformational families, we can compare them to established experimental evidence such as X-ray crystallography or high-field NMR spectra.

2.4.1 Epothilone A

As mentioned previously, our lab took an interest in the conformational preferences of epothilone A and its potential pharmacophoric similarity to paclitaxel. Though early experiments established the likely major solution conformer of epothilone A, we believed that its structure would allow for additional low-energy conformations in solution. As such, our lab opted to test the newly-developed polar map methodology on a conformational study of epothilone.\textsuperscript{145}

The molecular modeling analysis of epothilone A was performed using Schrödinger’s MacroModel software.\textsuperscript{196} A 10,000-step Monte Carlo (MCMM) conformational search was utilized along with the MM2* force field for minimization and the generalized Born/surface area (GB/SA) water model. Multiple searches were carried out using different starting geometries and ring-closure bond choices to fully explore the conformational space. After minimization, unique structures that fell within 30 kJ/mol of the global minimum were retained, resulting in a library of 187 conformations. Relative energies and backbone dihedral angles were extracted from each
of these using MacroModel’s *Filtr* command, which were subsequently plotted as polar coordinates using CricketGraph to yield a set of polar maps for conformational analysis.

Figure 2.3 – Epothilone A polar coordinate maps for C1-C9. Figure taken from Taylor, R.E.; Zajicek, J. *J. Org. Chem.* **1999**, *64*, 7224.

As seen in Figure 2.3, the initial report focused on six of the fifteen rotatable macrolide bonds, representing the C1-C9 segment of epothilone A. Five of the six exhibited two or three conformational preferences while the C5-C6 dihedral showed only one. Plotting these torsions against one another revealed that this region prefers to exist in two conformations, which were calculated to differ in steric energy by roughly 1 kcal/mol (Figure 2.4). Of particular interest was the lower energy conformer, which was similar to the previously-determined solid state conformation. The changes between these two conformations was proposed to be due to differences in the internal hydrogen bonding of the C3-hydroxyl (C5 ketone in Conformer A, C1 carbonyl in Conformer B) and the fact
that the C6 and C8 methyl groups allow for only two conformations that avoid high-energy syn-pentane interactions.

![Figure 2.4 - Solution conformations of epothilone A in CD$_2$Cl$_2$](image)

<table>
<thead>
<tr>
<th>no.</th>
<th>$^1$H shift (ppm, CD$_2$Cl$_2$)</th>
<th>NOE</th>
<th>atomic distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>conformer A</td>
</tr>
<tr>
<td>2a</td>
<td>2.40</td>
<td>2a–7 (m)</td>
<td>4.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2a–21 (m)</td>
<td>2.63</td>
</tr>
<tr>
<td>3</td>
<td>4.18</td>
<td>3–6 (s)</td>
<td>4.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3–7 (m)</td>
<td>6.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3–21 (s)</td>
<td>3.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3–22 (s)</td>
<td>2.68</td>
</tr>
<tr>
<td>6</td>
<td>3.21</td>
<td>6–9a (m)</td>
<td>2.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6–21 (s)</td>
<td>2.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6–24 (s)</td>
<td>4.77</td>
</tr>
<tr>
<td>7</td>
<td>3.73</td>
<td>7–23 (s)</td>
<td>2.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7–24 (s)</td>
<td>2.38</td>
</tr>
<tr>
<td>8</td>
<td>1.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9a</td>
<td>1.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>1.10</td>
<td>21–23 (s)</td>
<td>2.39</td>
</tr>
<tr>
<td>22</td>
<td>1.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>1.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>1.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.4 – Solution conformations of epothilone A in CD$_2$Cl$_2$ and selected chemical shifts, NOEs and calculated distances (s = Strong, m = Medium). Figures from Taylor, R.E.; Zajicek, J. J. *Org. Chem.* **1999**, *64*, 7224.

To verify the conformational analysis, NOESY and ROESY spectra were collected in CD$_2$Cl$_2$ (Figure 2.4). While most of the NOEs confirmed the presence of the major conformation (Conformer A), several anomalous NOEs were also observed. Specifically, the H2-H7, H3-H6, H3-H7 and H6-H24 could not have been present in Conformer A, supporting at least the minor presence of the alternative Conformer B.
These results spurred additional conformational studies, which led to the design of our lab’s conformational analogues of epothilone covered in Section 1.6.1.

2.4.2 Peloruside

![Figure 2.5 – The polyketide natural product peloruside A.](image)

Following the successful application of our polar map methodology to the epothilones, our lab sought to find new polyketides for conformational analysis. The next molecule in line turned out to be peloruside A, a polyketide natural product isolated from the New Zealand marine sponge *Mycale hentscheli* in 2000.\(^{197}\) This compound exhibited potent cytotoxicity with a mode of action similar to that of paclitaxel,\(^{198}\) and also possessed comparable activity in both normal and P-glycoprotein MDR-resistant cell lines.\(^{199}\) Interestingly, displacement studies with laulimalide found that peloruside shared its non-taxane binding site, and would also act synergistically with paclitaxel and epothilone.\(^{200}\) Low isolation yields and failed aquaculture attempts\(^{201}\) have left synthesis as the only means of obtaining this natural product, with several total syntheses reported to date.\(^{202-206}\) Although a number of analogues and congeners have been synthesized and evaluated, their severe loss of biological activity has shed little light on the overall SAR picture.
In 2006, Jiménez-Barbero and co-workers added their own piece to the puzzle by investigating peloruside’s free and bound solution conformations.\textsuperscript{207} For the solution conformation, they used conformational searches and molecular dynamics simulations to generate their conformational library, which they compared to D\textsubscript{2}O-derived NMR data. They found that peloruside A exists in two major conformations in solution, differing only in the relative orientation of the C9-C15 region (Figure 2.6). To determine the bioactive conformation they employed transfer-NOE experiments similar to those used by Carlomagno and Griesinger on the epothilones, and found that the TR-NOESY cross-peaks strongly supported solution conformer B as the bound conformation.

![Figure 2.6 – Overlay of peloruside major solution conformers A (blue) and B (green)](image)

The relatively small number of conformations observed by Jiménez-Barbero and co-workers in their MD simulations raised several questions in our laboratory, given that the lack of rigidifying features should not have biased conformational preferences to such a degree. Since their work only utilized the MM3 force field in their Monte Carlo conformational searches, our lab member Chris Nicholson decided to investigate the
effects of additional force fields as a way of fully exploring peloruside’s potential conformational space.\textsuperscript{208} By doing so, he hoped to open up possibilities for conformational analogues in a similar vein to our epothilone work.

![Polar maps of Peloruside A taken with the MM2* force field.](image)

Figure 2.7 – Polar maps of Peloruside A taken with the MM2* force field.

After employing the MM2, MM3 and OPLS-2001 force fields to generate conformational libraries, he used the previously described methodology to generate polar maps of peloruside’s backbone (Figure 2.7). As CricketGraph had become unsupported since the initial epothilone work, Nicholson took the opportunity to update the process of generating polar maps to use Microsoft Excel instead. These maps confirmed that the rigidity conferred by the C1 lactone ester (s-trans), C2,C3-diol monoether (gauche effect) and C5-C9 dihydropyran (chair) produced a relatively small number of conformational
families, while also maintaining the flexible C9-C15 region. Our lab hoped to gain additional insight into the relative populations of these conformational families by replacing the C11 and C13 stereocenters with a set of olefinic Z and E isomers. However, the completed peloruside analogues proved to be highly unstable, with the closed dihydropyran spontaneously equilibrating to the open keto-alcohol followed by subsequent degradation (Figure 2.8).^209

Figure 2.8 – Spontaneous equilibration of pyran lactol to open keto-alcohol

The bound conformation of peloruside A was not determined until 2014, with the publication of a tubulin-peloruside co-crystal structure that also confirmed its shared non-taxane β-tubulin binding site with laulimalide.^63 Interestingly, its bound conformation is highly similar to the solution structures derived by both our work and that of Jiménez-Barbero, differing only in the previously established C9-C15 region (Figure 2.9).
2.5 DISCON Analysis of Molecular Models

In the course of their conformational studies, the Smith lab needed a method to deconvolute their spongistatin solution NMR data into solution conformations. However, they found methods employing NMR constraints to be limiting, generating averaged conformations that tended to be higher in energy. They also concluded that purely computational methods derived from molecular mechanics force fields also suffered from inaccurate energy ordering of flexible molecules, due to the additive nature of the parameterization errors inherent to a given field. To circumvent these issues, they developed their own hybrid strategy to determine the most populated conformational families of a flexible organic molecule, the resulting program being termed DISCON (Distribution of Solution Conformations).

Though DISCON is effectively an implementation of NAMFIS methodology, the program offers several features that improve upon the latter’s shortcomings. One of the key issues with NAMFIS is that it attempts to fit every structure to a given data set
without regard to number of available NMR parameters. This is likely why Snyder’s first-pass analysis of laulimalide included several high energy structures, as their NMR pool was skewed towards higher-error NOE peaks versus coupling constants. DISCON avoids this with a routine to cluster conformations based on their theoretical NMR values prior to deconvolution. This allows the user to effectively group similar structures into clusters before fitting representative structures from each to the solution ensemble, which should avoid the “overfitting” problems of NAMFIS.

The DISCON software makes use of a wide variety of software packages. Janocchio\textsuperscript{210} is used for calculating theoretical NMR measurements from library structures, CLUTO\textsuperscript{211} is used for the clustering algorithm, and J Mol is used for 3D visualization and the graphical interface. Coupling constants are calculated using Altona’s modified version of the Karplus equation (which compensates for the electronegative effects of heteroatoms) when two sp\textsuperscript{3} stereocenters are present.\textsuperscript{186} Otherwise, the basic form of the Karplus equation is used. The software also includes the Balloon 3D Conformation generator\textsuperscript{212} in the event the user lacks the appropriate tools to generate a conformer library, though the software will easily accept output files from Schrodinger or other software platforms in the appropriate format.

2.5.1 Spongistatin

As mentioned in Section 1.5.7, the first application of DISCON came through the Smith lab’s conformational analysis of spongistatin.\textsuperscript{143} Their analysis began in similar fashion to those performed by our lab and the Snyder laboratory, using MacroModel searches to explore spongistatin’s conformational space. They employed the MMFF force field with a 100 kJ/mol window for their searches, which were performed not only with
the GB/SA solvation model for water but also DMSO, chloroform and acetonitrile. After elimination of redundant structures, their combined searches yielded 220 unique conformers for a final library. From the NMR side, they identified 11 $^3\text{J}_{\text{H-H}}$ coupling constants and 15 non-adjacent NOESY peaks for use with DISCON.

Figure 2.10 – Representative solution conformations of (+)-spongistatin 1 in DMSO. a) Twisted – 57% b) Flat – 13% c) Flat (CD ring flipped up) – 8% d) Flat (CD ring flipped down) – 4% e) Overlay of a-d with respect to the ABEF ring system.

For the DISCON analysis, they eventually settled at a clustering level of 25 structures. At this point, the ensemble generated lower violations to the observed NMR data when compared to the any single conformation. Overall, their study revealed four major conformational families that fulfilled the NMR constraints (Figure 2.10). The
major family, incorporating the low energy water conformation and termed “twisted”, contributed 57\% to the ensemble, while the other three “flat” families reflected different orientations of the CD ring in minor percentages (4-13\%). Separate calculations utilizing only coupling constants and only NOE constraints provided the major family with percentages of 53\% and 66\%, respectively, which signaled that their system had not been over-parameterized. Most importantly, the four conformational families shared the same linear orientation of the ABEF ring system along the western perimeter of spongistatin, which mirrored the results of their polar map analysis and led to the design of the simplified analogues outlined in Section 1.5.7.

2.6 Current Molecular Modeling Strategy

Our molecular modeling methodology has evolved over the years, but the same basic principles lie at the heart of our approach. In order to get a complete sampling of conformational preferences for later refinement, our model needs to initially predict as much of a molecule’s theoretical conformational space as possible. This is in sharp contrast to many other methods, which use NMR-based constraints to limit the scope of a conformational search. The complete set of predicted conformers are used to generate polar maps based on backbone dihedral angles, which are then used to determine conformational families that can be further refined with comparison to experimental data. Once the solution conformations have been obtained, analogues that take advantage of the overall conformational profile can be designed and biologically evaluated to determine the conformation-activity relationship. This section covers our current molecular modeling strategy, including modifications and improvements I have made to the process.
2.6.1 Conformational Search Methodology

The conformational analysis of a given compound begins with the *in silico* construction of its basic shape. Though there are a considerable number of programs available for this task, I have found that the MacroModel software package included as part of the Schrödinger Suite provides an intuitive interface and a wide variety of options. Molecules are built by first constructing the carbon backbone before installing additional functionalities and heteroatoms, periodically minimizing the entire structure to provide a decent starting geometry for the conformational search.

Choosing a force field is an important part of our methodology, as the parameters used will essentially determine the outcome of a conformational search. The most common force fields for simple organic molecules are part of the MM series developed by Allinger, particularly the MM2 and MM3 force fields that we used with epothilone and peloruside. However, these force fields were initially developed for hydrocarbon systems,\(^{191}\) which differ considerably from highly oxygenated species such as polyketide natural products. The iterations of these fields developed by Schrödinger (MM2* and MM3*) alleviate their limitations to a degree by incorporating additional parameters for features like conjugation and hydrogen bonding and improved treatments for electrostatics. However, Snyder has elaborated that several force fields should be used instead of relying on only one, given the variable treatment of polarization effects and electrostatics.\(^{213}\) Nicholson’s work on peloruside A found that use of the MM2, MM3, and OPLS-2001 force fields would each provide different conformational families and with unique global minimum conformations, which differed noticeably in potential energy when evaluated with B3LYP/6-31G DFT calculations (Figure 2.11).\(^{208}\) After
experimentation, our current methodology uses the MM3*, MMFF (Merck Molecular Force Field, based on MM3), and OPLS-2005 force fields to provide a good mixture of Class 1 and Class 2 fields.

Along with a given force field, we make use of the GB/SA implicit solvation model for water. This model provides a computationally-friendly mimic of the biological environment in which these molecules would be active. Additionally, conformational searches performed on larger macrocycles in vacuum or chloroform solvation models can experience hydrophilic collapse as a result of internal hydrogen bonding, which could result in otherwise unfavorable conformations.

Conformers are generated with MacroModel’s Batchmin conformational search module, using the Monte Carlo (MCMM) search algorithm. MacroModel defines two types of parameters, Ring Closures and Torsions, which it can freely manipulate to generate new conformations. As many polyketides will lack the two adjacent achiral atoms recommended for a ring closure bond, our methodology uses a C-O lactone bond if available. Ester functionalities preferentially adopt an \(s\text{-}cis\) configuration\(^{214}\) that we have
found to be exceptionally rigid from a torsional standpoint, which means the C-O bond should return to a 180° dihedral angle following manipulation. For allowable torsions, all backbone ring bonds are chosen with the exception of ring closure bonds. Other bonds, particularly substituent methyl, methoxy or hydroxy groups and longer side chains, are ignored to maximize macrolide conformational sampling and eliminate structures that would otherwise share identical ring conformations.

Our previous conformational analyses have performed searches ranging from 10,000 to 25,000 iterations (generated conformers). Experimentation on my part has set the current standard at 50,000 steps. I have generally found the number of new conformations begins to approach zero before this level, suggesting that the search has approached convergence. Each generated conformer is minimized using the default Polak-Ribières conjugate gradient method (5000 steps, convergence threshold of 0.05), and all unique energy-minimized structures falling within 20 kJ/mol (4.78 kcal/mol) of the instantaneous global minimum are saved for analysis. This energy window is used to provide for a thorough sampling of feasible conformations while keeping a reasonable number of overall structures for evaluation. Computational experiments by Nicholson have found that extending the energy window to 50 kJ/mol (11.95 kcal/mol) has a minimal effect on observed conformational families.208

In keeping with the multiple force field philosophy, separate conformational searches using each of our chosen force fields are run to yield individual libraries. These are then combined into a single large library and subjected to MacroModel’s Redundant Conformer Elimination routine, which removes structures based on the RMS similarity of selected atoms or torsions. I have found that the default settings of 0.5Å atomic RMSD
works best to eliminate redundant structures, using the backbone ring atoms for comparison. This process yields a final library suitable for further analysis, usually containing 200-400 unique conformations.

2.6.2 Graphical Clustering Analysis

The amount of data generated by a conformational search makes refinement and graphical representation difficult. Traditionally, refinement is accomplished with clustering algorithms based on RMS comparisons similar to those used for redundant conformer elimination, which will group similar structures. Previously our lab used the Superposition and XCluster modules, though these have now been superseded by the Clustering of Conformers cheminformatics script in the Schrödinger Suite (Figure 2.12).

![Figure 2.12 – Clustering of Conformers and Clustering Statistics windows for a sample library input](image)
This routine offers the ability to select either atomic positions or torsions for difference variables and uses them to calculate an RMSD matrix, which is subsequently clustered with the user’s choice of similarity settings. The results of this process are displayed by sorting the library into its constituent clusters based on user input, though the script itself has a suggested level based on the Kelley index criterion. This technique provides for a graphical means of directly comparing clustered compounds to one another. Currently, our strategy is to use the backbone atoms in order to generate the RMSD matrix, which serves as a useful way to provide a preliminary visualization of potential conformational families. If areas of the molecule prove to be particularly flexible regardless of other local conformational preferences, these atoms can be omitted to cluster around conformationally rigid areas instead.

2.6.3 Polar Map Generation

Though preliminary evaluation of a conformational library via clustering can be useful, information on how conformations change as a function of potential energy is not present with this analysis, particularly with respect to individual conformational changes produced by local acyclic interactions. As such, a thorough analysis of a molecule’s conformation with our methodology is accompanied by a full set of backbone polar maps, which will allow us to plot dihedral angle changes against a structure’s force field-derived potential energy. This strategy relies on the assumption that many of the unique structures derived from the conformational search will be related to one another, and that viewing a molecule as a series of individual rotatable bonds will enable the visualization of the entire library as a series of small, related fragments.
For our analysis, we select a desired dihedral angle in Maestro, the graphical interface used to launch MacroModel, which generates a torsional value that is associated with each selected conformation. After repeating this process on each dihedral, the conformational data is extracted and tabulated into a spreadsheet. To generate the maps, the data has to be plotted as polar coordinates \((r, \theta)\), where \(r\), the radius term, is the energy relative to that of the global minimum conformer (which represents the origin) and \(\theta\) represents the dihedral angle (Figure 2.13).

![Figure 2.13 – Sample polar coordinate map and analysis of a dactyloolide torsional angle (see Section 4.4.1 for full analysis). Each green dot on the plot indicates a single conformer.](image)

The initial methodology made use of CricketGraph as a means of easily generating the needed polar maps, but the demise of this program has shifted our process to use Excel in order to increase longevity. While converting the conformational data to polar coordinates through trigonometric means is possible, in practice this involved
dealing with multiple templates since Excel lacks the native functionality to create these maps. However, we discovered a Polar Plot add-in to Excel that generates polar coordinate graphs directly from the conformational data without manipulation, which in practice greatly accelerated the process of creating a full set of polar maps. Additionally, this allows for the easy creation of Boltzmann distributions if desired, by calculating the population values of a set of conformers based on their potential energy relative to the global minimum and feeding the data back into the add-in.

2.6.4 DISCON Analysis

Though our methodology still uses polar maps at the heart of its strategy, the appearance of DISCON intrigued me. The problems with NAMFIS elaborated previously prevented us from employing it in our own analyses, but the hierarchical clustering used by DISCON presented a softer alternative to identifying solution conformations. Additionally, while polar maps are effective at identifying local conformational preferences, extremely large and flexible molecules (i.e. spongistatin) complicate the identification of conformational families due to the sizeable output of a conformational search. Since the identification of solution conformations through comparison to experimental data is the ultimate goal, I have incorporated DISCON into our methodology.

Use of DISCON requires both the conformational library generated earlier and a full panel of 1D and 2D NMR data. From this data, a complete or near-complete set of coupling constants is needed along with a set of NOEs, both of which require the unambiguous identification of every proton on the structure. The coupling constants are used to define torsional angles while the NOE cross peaks are used to define long range
interactions, which means that the data set used for DISCON analysis needs to be carefully selected such that terms do not overlap with one another, which could cause overfitting. DISCON uses quantitative NOE cross peaks as a means of determining interproton distances, which it normalizes based on a user-defined reference cross peak between two protons whose distances will not change between conformations (i.e. two germinal methylene protons or two adjacent aromatic hydrogens). While the program will accept NOE peak volumes, I have found that calculating experimental interproton distances can make for a better visual aid and potentially alleviate errors with DISCON’s internal calculation. For ROESY cross peaks, interproton distances are calculated with the equation series below:  \[ r_{ij} = r_{ref} \left( \frac{a_{ref} c_{ref}}{a_{ij} c_{ij}} \right)^{1/6} \]

where

\[ c_{ij} = \left( \frac{1}{\sin^2 q_i \sin^2 q_j} \right) \]

and

\[ \tan q_i = \left( \frac{g B_1}{w_1 - w_0} \right) \]

In this series, \( a_{ij} \) is the NOE cross-peak volume, \( r_{ij} \) is the interproton distance between protons i and j, \( (w_1 - w_0) \) is the difference in Hz between the chemical shift of the cross-peak and the transmitter center (o1p), and \( gB_1 \) is the spin lock power. Given a known distance between two protons \( (r_{ref}) \) and the resulting NOE volume \( (a_{ref}) \), the interproton distance between a cross peak between protons i and j can be calculated.
Because ROESY cross-peak intensities are less intense the further they are from the transmitter center, the additional ‘c’ term provides the necessary offset.

Once the data has been entered, the DISCON calculation can be run. Observations from the initial run are useful for determining the suitability of the chosen NMR parameters, as the sample dendogram shows in Figure 2.14. Here, NMR parameters are color coded such that the lowest value will be shown in red and the largest will be shown in green. If a variable is colored the same across all conformations, the parameter is essentially unchanged across the conformational space, and can be safely removed in subsequent runs.

Figure 2.14 – Sample DISCON dendogram from http://discon.sourceforge.net. Columns of identical or nearly identical color throughout may be safely omitted in future runs.
One of the important considerations with DISCON is that multiple runs are necessary to determine the appropriate clustering level and identify the representative conformational families. Initially, multiple cluster levels are tested, typically in multiples of five (5, 10, 15, etc.), and both the RMS deviation and error level of the solutions are examined. Ideally, the RMSD should be lower than that of any single conformation, and the ideal cluster level should be where the solution error has achieved a minimum. Through observation of the different randomized solutions at the desired cluster level, representative conformers from each cluster that can be identified for use in further analysis.

2.6.5 Final Visual Analysis

Though DISCON purports to provide a quantitative answer for solution conformational families, the software is largely a “black box”. DISCON is mostly undocumented and the combination of several algorithms (clustering, calculation of theoretical NMR values, etc.) means that using it for the final word on a conformational analysis would be foolish. As such, I have found that it is critical to finish with a visual inspection of its suggested conformers and their subsequent comparison to experimental NMR data to yield the final representative solution conformations.
3.1 Purpose

The purpose of this chapter is to introduce the polyketide natural product neopeltolide, related synthetic studies, and our efforts towards identifying its solution conformation and binding site. The isolation, biological activity and purported mode of action will be discussed, followed by a summation of previous synthetic work and structure-activity relationship studies. The conformational analysis of neopeltolide’s macrolide core will be presented with similar analyses of known analogues to illustrate how they alter conformational preferences, along with a computational analysis of C2-substituted analogues to illustrate how they may not. Separate computational studies will be covered to highlight a potentially novel binding site on neopeltolide’s purported biological target, along with our in silico attempts to identify the true binding site of the polyketide. Finally, the total synthesis of neopeltolide and a C2-methyl analogue will be presented.

3.2 Background

3.2.1 Isolation and Biological Activity

In 2007, Wright and co-workers reported the isolation of a new natural product, neopeltolide, from a deep-water sponge of the family Neopeltidae, genus Daedalopelta,
A 105 g sample of the sponge was exhaustively extracted with ethanol, and the resulting crude residue was separated by vacuum-column chromatography to yield an enriched fraction. Further separation of this fraction using reverse-phase HPLC yielded a small amount of purified neopeltolide (4.0 x 10^{-3} % of frozen weight) as colorless oil.

Despite the small amount of material available, Wright and co-workers were able to perform several assays to gauge biological activity. Neopeltolide was found to be a potent inhibitor towards the growth of the fungal pathogen *Candida albicans*, which can severely affect patients with immunocompromised systems. Additionally, neopeltolide was found to be an extremely potent inhibitor of tumor cell proliferation *in vitro*, with cytotoxicity assays revealing nanomolar concentration IC_{50} values of 1.2, 5.1 and 0.56 for the A549 (lung), NCI/ADR-RES (ovarian), and P388 (leukemia) cell lines, respectively. Neopeltolide also demonstrated strong inhibitory effects towards the PANC-1 and DLD-1 cell lines, though the observed 50% cell death over an extended dose range suggests that it is cytostatic towards these lines instead of cytotoxic. Later work performed by the Floreancig group showed that neopeltolide is considerably less active towards a variant of the MCF-7 breast cancer cell line, suggesting that it and its analogues demonstrate cell line selectivity rather than possessing general cytotoxicity.\textsuperscript{218}

3.2.2 Structural Determination and Stereochemical Reassignment

Following isolation, the molecular formula and structure of neopeltolide were proposed by the Wright lab on the basis of their spectroscopic analysis.\textsuperscript{217} They determined that neopeltolide contains a 14-membered macrolactone containing a 2,4,6-trisubstituted tetrahydropyran, with six stereocenters in total. Additionally, the structure
was found to possess an oxazole-containing side chain identical to that of leucascandrolide A (3.2, Figure 3.3), a previously-discovered polyketide with notable antiproliferative activity.\textsuperscript{219}

Figure 3.1 – The structure of neopeltolide as originally reported by Wright and co-workers and the revised structure reported by the Scheidt and Panek laboratories. Key NOE interactions for 3.1a are highlighted in red.

The assignment of neopeltolide’s stereocenters was based on the analysis of coupling constants, a 2D-NOESY spectrum and a series of 1D-DPFGSE NOE experiments. The pyran protons H3 and H7 were assigned as axial based on their strong NOE interaction, while H5 was assigned as equatorial due to smaller coupling constants and a lack of cross-peaks with either H3 or H7. The relative stereochemistry of the C3, C7, C9, C11 and C13 stereocenters was assigned on the basis of observed NOE enhancements between H3-H7, H7-H9, H9-H11, and H11-H13, with all protons on the
same side of the macrolide ring (Figure 3.1). Unfortunately, because of a lack of material Wright and co-workers were unable to obtain a crystal structure to determine the absolute stereochemistry.

Although the Wright group was able to tentatively assign the relative stereochemistry of neopeltolide, the absolute stereochemistry would have to be determined via chemical synthesis. Due to its promising biological profile and intriguing structure, neopeltolide attracted considerable attention from the synthetic community. However, two early syntheses of 3.1a, performed independently by Panek et al.\textsuperscript{220} and Scheidt et al.,\textsuperscript{221} found that the reported structure of neopeltolide did not match the published spectroscopic data. After ruling out synthetic discrepancies, both groups proposed that the correct stereochemistry of neopeltolide was 3.1b, confirming this assignment with independent total syntheses.

In his thesis, Youngsaye provided an explanation for the initial error in structural reassignment.\textsuperscript{222} The proton spectrum of neopeltolide possesses considerable overlap in the aliphatic region of its proton spectrum in $d_4$-methanol, which includes several of the key peaks used to determine NOE enhancements. The C9 proton peak ($\delta = 1.38$ ppm) closely overlaps with the protons from the neighboring C8 methylene ($\delta = 1.36$ ppm) and C12 methylene ($\delta = 1.28$ ppm) groups, which cast doubts on the H7-H9 and H9-H11 NOE peaks. As these could now conceivably represent H7-H8 and H11-H12 peaks, respectively, they concluded that they had no information as to the identity of the C9 stereocenter. As Panek and co-workers were confident that the configuration of the tetrahydropyran ring and the C11, C13-\textit{syn} assignment were both correct, they
rationalized that there were four possible configurations of the C9, C11 and C13 stereocenters, one of which would be the absolute stereochemistry.

As the original Wright structure represents one of these stereoisomers, the Panek group prepared the remaining three synthetically. Upon comparison of the compounds to the isolation spectra, they concluded that isomer 3.1b was the correct structure of (+)-neopeltolide.

3.2.3 Formal and Total Syntheses of Neopeltolide

The successful identification of neopeltolide’s absolute stereochemistry led to a significant increase in attention from the synthetic community. Panek’s\textsuperscript{220} and Scheidt’s\textsuperscript{221} total syntheses were followed by reports from the Lee,\textsuperscript{223} Kozmin,\textsuperscript{224} Paterson,\textsuperscript{225} Fuwa,\textsuperscript{226-228} Roulland,\textsuperscript{229} A. Ghosh\textsuperscript{230} and Hoveyda\textsuperscript{231} laboratories, along with a number of formal syntheses from the Maier,\textsuperscript{232} Taylor,\textsuperscript{233} Floreancig,\textsuperscript{234} Hong,\textsuperscript{235} Jennings,\textsuperscript{236} Yadav,\textsuperscript{237} She,\textsuperscript{238} Sharma,\textsuperscript{239} S. Ghosh\textsuperscript{240} and Raghavan\textsuperscript{241} laboratories. The vast majority of this work was focused on the synthesis of neopeltolide’s macrolide core, with total syntheses using methodology developed by one of the Kozmin,\textsuperscript{242} Leighton\textsuperscript{243}
or Panek\textsuperscript{244} groups for leucascandrolide A to furnish the oxazole-containing side chain. These synthetic routes have been exhaustively reviewed and contrasted elsewhere.\textsuperscript{245}

3.2.4 Kozmin: Neopeltolide’s Mode of Action

As mentioned earlier, one of the notable features of neopeltolide is its strong resemblance to leucascandrolide A, possessing not only the same oxazole-containing side chain but also sharing the same 2,4,6-tetrahydropyran subunit. Given their status as both antifungal and antiproliferative agents it reasonably follows that they would share similar biological targets, and the Kozmin group formed the hypothesis that neopeltolide might be considered a simplified analogue of leucascandrolide A.\textsuperscript{224} As such, they began preliminary investigations in 2008 into the biological target and mechanism of action for both compounds.

Figure 3.3 – Screening compounds used by Kozmin and co-workers in their mode of action study. Areas of change on leucascandrolide A are highlighted in orange.
Kozmin and co-workers began their study with an analysis of leucascandrolide A. They noticed that its non-natural enantiomer was only two- to three-fold less potent than the natural product,\textsuperscript{246} which led them to explore simplifying the macrocycle. Following conformational analysis, they settled on a structure that removed the C12 methyl, C18-C19 alkene and C21 methyl groups, yielding the conformationally-similar analogue 3.3. Synthesis of this analogue proceeded smoothly, achieving the finished product in 24 total steps (as opposed to the 34 total steps needed for 3.2)\textsuperscript{242,246}. Fortuitously, this analogue showed biological activity equal to that of leucascandrolide A against A549, PC3 and HCT116 tumor cell lines. They also reported their own route to 3.1b, which was delivered with a longest linear sequence of 15 steps. Evaluation of the effect of these compounds on cell cycle regulation was also fruitful, as all three showed the same G1 cell cycle delay that was reported for neopeltolide.\textsuperscript{217}

After establishing that their compounds would be useful as enabling tools, Kozmin and co-workers began their search for the shared mechanism of action. They screened 3.3 against a library of 4900 yeast strains with different gene deletions,\textsuperscript{247} which quickly highlighted the SNF4 gene as important. This gene encodes a regulatory subunit of the yeast homolog of the stress-responsive mammalian AMPK, which is normally triggered by inhibition of ATP production.\textsuperscript{248} The sensitivity of the SNF4 deletion mutant to 3.3 indicated that the compound might be targeting mitochondrial oxidative phosphorylation. Follow-up experiments identified the mitochondrial electron transport chain as the site of action for 3.3 and neopeltolide, and experiments with purified enzymes finally established cytochrome bc\textsubscript{1} (complex III) as a target of both compounds. Interestingly, neither compound was able to inhibit the other complexes of the electron
transport chain, despite the fact that they also utilized the same cofactor (coenzyme Q10) as cytochrome $b_{c_1}$.

Figure 3.4 – Some of the known inhibitors of cytochrome $b_{c_1}$. The two forms of coenzyme Q10 (ubiquinol and ubiquinone) are boxed at the bottom.
Cancerous cells are particularly susceptible to disruption of oxidative phosphorylation due to their increased dependency on glycolysis, known as the Warburg effect.\textsuperscript{249-251} However, while cytochrome $bc_1$ is a fairly well-established target for inhibitors of several different pathogens, its significance in cancer cell proliferation is much less clear.\textsuperscript{224,252} Additionally, known inhibitors of this complex share little structural similarity with neopeltolide or leucascandrolide A, and it is uncertain as to whether the observed inhibition of cytochrome $bc_1$ fully accounts for the antiproliferative effects of these compounds.

3.2.5 Analogue Work and SAR Studies

Alongside the reports of total and formal syntheses, several groups have utilized the various synthetic routes to perform thorough structure-activity relationship studies on the various components of neopeltolide. Despite its structural complexity, neopeltolide has proven to be an astonishingly flexible scaffold for modification, with analogues possessing altered macrolides almost universally retaining biological activity against a range of cancer cell lines. Conversely, changes to the oxazole-containing side chain generally result in a severe loss of activity.

3.2.5.1 Side Chain Analogues

The necessity of the neopeltolide side chain was first explored by Scheidt and co-workers.\textsuperscript{253} Studies performed on leucascandrolide A established that the side chain \textit{3.4} on its own possesses fungistatic activity,\textsuperscript{219} but the Scheidt group found that it had no anticancer activity when separated from the macrolide. Additionally, the core pyran-alcohol (\textit{3.5}) and core pyranone (\textit{3.6}) possessed only nominal inhibition. Attempts to
replace the side chain with either benzoyl (3.7) or octanoyl (3.8) esters resulted in a
dramatic loss of potency against MCF-7 cells (5.8 ± 4.7 µM and >10 µM, respectively,
vs. 2.2 nM for 3.1b).

![Chemical structures](image)

Figure 3.5 – Scheidt’s analogues of neopeltolide.

Maier and co-workers would expand upon this SAR by investigating the tail
connecting the oxazole/carbamate portion of the side chain to the macrolide, concluding
that changing the olefin geometry (3.9) or shortening the carbon chain (3.10, 3.11) leads
to a significant loss of potency. Additionally, rigidifying certain areas of the alkyl
chain via olefination also has subtle effects on biological activity, with the Z,E-dienoate
3.12 having the closest activity to natural neopeltolide while the E,E-dienoate 3.13 saw a
ten-fold decrease.
Owing to the difficulties in preparing the oxazole subunit and the Z-allylic carbamate, Floreancig and co-workers next investigated whether side chains that contained different aromatic functionalities would show comparable biological activity. Accordingly, they synthesized several side chains with a pyridine (3.14), a \textit{meta}-substituted benzene (3.15), and a furan (3.16) in place of the oxazole ring. To determine the necessity of the carbamate and the ester-linkage, they also synthesized des-carbamoyl (3.17) and amide-linked (3.18) side chains, respectively. In general, these alterations to the neopeltolide side chain resulted in noticeable decreases in biological activity, with the furan-containing analogue showing the most potency.
The most recent side chain analogues were reported by Fuwa and co-workers, who effectively completed the SAR studies.\textsuperscript{255} Their work investigated alterations to the C19-C20 olefin via both saturation and substitution, and used Floreancig’s 8,9-dehydroneopeltolide macrolide\textsuperscript{234} as the base. They also investigated the role of the C26-C27 olefin separately from the carbamate tail, as the Floreancig group had omitted it in analogue \textit{3.17}. Fuwa and co-workers determined that these analogues generally lost biological activity, with the notable exception of C19- and C20-methyl compounds \textit{3.20} and \textit{3.21} (IC\textsubscript{50}: 2.5 nM and 6.3 nM, A549 cells). As saturation of the C19-C20 bond resulted in a loss of potency, they suggested that it could act as a conjugate acceptor.
3.2.5.2 Macrolide Analogues

Early macrolide analogues of neopeltolide were generated using fragments or diastereomers made during the course of total synthetic work. Maier and co-workers used the 5-epi macrolide minor diastereomer from their total synthesis to generate 5-epi-neopeltolide 3.28, and used the enantiomer of a chiral allylation reagent in an earlier synthetic step to generate 11-epi-neopeltolide 3.29 using the remainder of their synthetic route.\textsuperscript{254} During the course of their own total synthesis, the Scheidt lab generated two
diastereomers of the macrolide in their efforts to elucidate the correct absolute stereochemistry of neopeltolide and test their synthetic route for discrepancies, including Wright’s originally-proposed macrolide and its C5 epimer. They subsequently connected the oxazole side chain to these macrolides to provide full neopeltolide analogues 3.1a and 3.30 for their biological studies.253 As shown in Figure 3.9, all of these macrolide analogues retained nanomolar activity, though less than that of neopeltolide 3.1b.

![Chemical structures of Maier and Scheidt macrolide analogues of neopeltolide](image)

**Figure 3.9** – Maier and Scheidt macrolide analogues of neopeltolide. The oxazole side chain (R) has been omitted for clarity.

The Floreancig group took a different approach to their neopeltolide analogues by employing late-stage diversification to explore the C8-C9 region of the macrolide. As their route took them through an 8,9-dehydro macrolide intermediate (3.31) they took the opportunity to selectively hydrogenate and oxidize the C8-C9 olefin to yield six different
analogue, including the aforementioned 3.19.\textsuperscript{218,256} As shown in Figure 3.10, these analogues all retained significant biological activity save for the 8,9-dihydroxy analogue 3.34, showing that changes to the C8 and C9 positions are reasonably well-tolerated.

![](image)

Figure 3.10 – Floreancig macrolide analogues, based on 8,9-dehydro macrolide 3.31. The oxazole side chain (R) has been omitted for clarity. All analogues were tested against HCT116 cells, with neopeltolide (3.1b) possessing Gl\textsubscript{50}: 0.77 nM.
In addition to their side chain analogues, Fuwa and co-workers have performed the most thorough investigation of neopeltolide’s macrolide SAR to date. Along with their total synthesis report, they used their route with two early stage intermediates to yield 9-desmethyl (3.36) and 11-desmethoxy (3.37) neopeltolide, both of which retained low nanomolar activity. Additionally, 3.36 proved to be even more potent than the parent compound in P388 murine cells (IC\textsubscript{50} of 0.813 nM vs. 0.899 nM for 3.1b). Their later work made use of Floreancig’s 8,9-dehydro scaffold, in which they systematically generated fifteen new diastereomeric analogues to comprehensively cover the biological role of each stereocenter, as seen in Figure 3.11. Along with this impressive library they also generated seven structurally-simplified analogues of neopeltolide. Two of these investigated the level of substituent removal that the macrolide would tolerate, and the remaining five as to whether the macrolide is necessary at all for biological activity.

Figure 3.11 – Fuwa’s neopeltolide analogues. The oxazole side chain (R) has been omitted for clarity. All IC\textsubscript{50} values were taken against A549 cells unless indicated otherwise.
Fuwa and co-workers found that while modifications to the stereochemistry of the C11 and C13 carbons were generally well-tolerated, the presence of the C13 \( n \)-propyl tail is necessary for low-nanomolar inhibition. Additionally, the relative stereochemistry found in the macrocyclic pyran must be maintained, as changing the C5 side chain linkage from axial to equatorial was met with a significant loss of biological activity. Interestingly, the Maier group found that their 5-\( epi \) analogue \textit{3.28} possessed an IC\textsubscript{50} of 5.6 nM against A549 cells\textsuperscript{254} while Fuwa’s similar 5-\( epi \)-8,9-dehydro \textit{3.19h} dropped all the way to 2.8 \( \mu \)M against the same cell line.\textsuperscript{257} Of the greatest interest were their structurally-simplified analogues, which retained nanomolar inhibition of A549 cells. Even more peculiar were the non-macrocyclic analogues, one of which (\textit{3.42}) possessed an IC\textsubscript{50} of 43.9 nM against A549 cells despite lacking most recognizable functionalities of neopeltolide.

3.2.5.3 Structural Conclusions

![Overall SAR of neopeltolide](image-url)

Through compilation of the different analogue studies, Fuwa and co-workers summarized the overall SAR profile of neopeltolide in Figure 3.12. The oxazole-containing side chain is essential for activity and must be axially attached to the tetrahydropyran ring, which in turn must retain the absolute configuration of neopeltolide. The remainder of the macrolide serves as conformational control to properly orient the individual substituents, particularly the C13 \( n \)-propyl tail, to achieve nanomolar inhibition. However, the substituents themselves are highly amenable to modification, though doing so will sacrifice the maximum possible inhibition level observed in the parent natural product. Of note is the fact that a large percentage of this SAR was performed in A549 cells rather than across multiple cell lines, and as neopeltolide does not appear to possess general cytotoxicity\(^{218} \) it is possible that additional factors may be at play.

3.3 Polar Map Analysis of Neopeltolide

The stereochemical misassignment of neopeltolide in Wright’s initial report raises some interesting questions about its solution conformational preferences. NMR data observed by both the Scheidt and Wright laboratories strongly suggests that neopeltolide’s macrolide should possess multiple conformations, as both the H9-H11 and H11-H13 cross peaks do not appear to be likely in any single conformation with the revised stereochemistry. Preliminary conformational searches performed in our lab found two low energy conformations that could explain the observed NMR data, which appeared to be controlled by the movement of the \( s-cis \) ester linkage (Figure 3.13). As a result, our lab embarked on a program to determine neopeltolide’s solution
conformational preferences to guide the design of analogues that would share its impressive biological activity.

![Diagram of neopeltolide conformations](image)

Figure 3.13 – Preliminary modelling work showing the proposed conformations that would account for the observed H9-H11 (green) and H11-H13 (red) NOE crosspeaks in the correct stereochemical assignment of neopeltolide.

3.3.1 Analysis of Neopeltolide

As the SAR studies of neopeltolide performed by Fuwa et al have demonstrated, the conformation of the neopeltolide macrolide appears to be responsible for the difference between mid- and low-nanomolar antiproliferative activities. As such, we decided that including the oxazole-containing side chain as part of a conformational search would be inefficient, since it would add flexible torsions that
would not contribute to macrolide conformation-activity relationship hypotheses. Modeling of the neopeltolide macrolide was performed using the methodology detailed in Section 2.6.1. Briefly, conformational searches were performed using Macromodel 9.9 with 50,000-step Monte Carlo searches using the MM3* force field along with the GB/SA solvation model for water, saving conformers that fell within a 20 kJ/mol window of the global minimum. Each conformer was minimization-converged and found multiple times, suggesting that the search was convergent. Results from the search were further pruned through redundant conformer elimination to yield the final library. Conformational searches employing other initial starting geometries and the MMFF* and OPLS2005 force fields were also performed, though these did not reveal any new data and further suggests a complete exploration of the conformational space.

Figure 3.14 – Polar maps of the neopeltolide macrolide.
Once the library was established, polar maps were generated using the methods detailed earlier (Figure 3.14). Interestingly, the maps revealed that the macrolide does not appear to possess a great deal of flexibility, with most of its torsions adopting only one or two possible conformations. As expected, the $s$-$cis$ ester was able to rotate freely, but curiously appeared to do so with little effect on the remainder of the macrolide. By clustering the library structures, we found that many of the conformational families would possess a nearly identical conformation with the ester flipped 180°, differing in potential energy by less than 1 kJ/mol (Figure 3.15). The western half (C7-C11) was responsible for the bulk of neopeltolide’s flexibility, though the polar maps indicated that conformations adopting these different dihedral angles were noticeably (>4 kJ/mol) higher in energy than the global minimum.

Figure 3.15 – Overlays of two conformational families showing the difference in ester conformation. The two conformers represented in each family differ by >1 kJ/mol (MM3*).
Following this initial analysis, we took the opportunity to evaluate neopeltolide with solvation models other than that of water. As mentioned previously, our methodology traditionally uses the GB/SA water solvation model to avoid hydrophilic collapse of the macrocycle due to internal hydrogen bonding. However, prior work by our lab had established a nearly full set of coupling constants in CDCl$_3$, and we hypothesized that the smaller neopeltolide macrolide would be extremely unlikely to collapse due to ring strain. Gratifyingly, repeating our methodology using the solvation model for chloroform resulted in virtually identical polar maps to those produced by the water model, suggesting that neopeltolide’s macrocyclic conformational preferences should be unchanged in organic solvents.

3.3.2 Analysis of C2-Substituted Macrolide Analogues

![2-(S)-macrolide and 2-(R)-macrolide](image)

Figure 3.16 – Expected conformational results of C2 methyl substitution on the macrolide s-cis ester.

As mentioned previously, preliminary conformational work on neopeltolide suggested that macrocyclic conformation was subject to the orientation of the s-cis ester. We hypothesized that addition of steric bulk to the C2 position could attenuate the observed free rotation of this functional group through allylic strain, similar to what we
observed in our work with C14-substituted epothilone analogues. Our hope was that the (R)-2-methyl and (S)-2-methyl macrocyclic epimers would possess different conformational preferences (Figure 3.16). Accordingly, both analogues were built in MacroModel and subjected to the same methodology as before.

Unfortunately, the resulting polar maps failed to reflect the desired changes in the macrocycle’s conformational profile. In the epothilones, installation of the C14 methyl group created allylic strain that resulted in the complete elimination of the disfavored structure from the conformer library. Conversely, the polar maps from the neopeltolide 2-methyl epimers were virtually unchanged from those of the natural macrolide (Figure 3.17). Increasing the size and constitution of the group installed at the C2 position largely

Figure 3.17 – Comparison of the (R)-2-methyl and (S)-2-methyl macrocyclic epimer polar maps.
had little effect, with allyl, isopropyl and trifluoromethyl groups demonstrating similar conformational preferences in the O-1-2-3 torsion. Changing the group to either a hydroxyl or methoxy did finally achieve a different polar map for this dihedral angle, removing a wide sweep of potential conformations (Figure 3.18).

![Figure 3.18 – Comparison of different C2-substituted macrolide O-1-2-3 polar maps.](image)

Normally, alkyl substitution α to a ketone should cause the new group to adopt an eclipsed conformation to the carbonyl. This preference will keep the alkyl groups on either side of the ketone *anti* to one another, which should avoid steric interactions. However, with an ester this preference should disappear as both oxygens will have roughly the same steric properties. Conversely, introducing an electronegative element (like a hydroxyl group) could potentially influence the ester conformation through an
anomeric effect. The anomeric effect describes the tendency of heteroatomic substituents adjacent to a heteroatom in a cyclohexane ring to adopt an axial conformation over an equatorial one due to $\sigma^*$ hyperconjugation,\textsuperscript{260} but Denmark and co-workers also found that this effect could occur with $\pi$-$\sigma^*$ hyperconjugation in cyclohexanone rings.\textsuperscript{261} In our case, this would bias the O-1-2-3 torsion to adopt either the 100° or 0° regions. It is also possible that the observed effects could be due to internal hydrogen bonding between the C2 hydroxyl and the lactone carbonyl, an effect which is further explored in Section 5.4.2.4.

![Figure 3.19 – The two possible ester orientations of Conformer A for (R)-2-methyl neopeltolide.](image)

3.3.3 Analysis of Previous Neopeltolide Analogues

As mentioned previously, several analogue studies of neopeltolide focused on alterations to the macrolide. Since many of these compounds retained considerable potency, we wanted to see if the macrocyclic conformation of these analogues was a factor in their biological activities. To gauge each of the major macrocyclic substituents, we selected the 9-\textit{epi} (3.35),\textsuperscript{256} 11-\textit{epi} (3.29),\textsuperscript{254} and 11,13-\textit{epi} (3.1a)\textsuperscript{253} analogues that were generated by Floreancig, Maier and Scheidt, respectively. We also looked at the
8,9-dihydroxy analogue 3.34 generated by Floreancig, as this represents the only macrocyclic analogue to date to drastically lose biological activity.²¹⁸

Each analogue was built, searched and mapped using the same methodology as our other structures. As one would expect, each structure departs significantly from the natural macrolide in terms of conformational preferences, with much of the deviation occurring near the change in structure. The 9-epi analogue is mostly similar to the natural structure but shows notable changes from C7-C11. The 7-8-9-10 map shows a loss of structures from 240-280° while the formerly high energy structures with torsional angles of 60° become far more energetically feasible. Conversely, the 8-9-10-11 and 9-10-11-12 polar maps exclusively show losses of conformations at 70° and 300°, respectively. These changes are significant enough for the 9-epi global minimum structure to

Figure 3.20 –Backbone polar maps and global minimum conformation for 9-epi neopeltolide 3.35.
noticeably depart from neopeltolide’s, as seen in Figure 3.20, which appears to be a result of the molecule attempting to keep the C9 methyl group in a pseudo-equatorial position.

Figure 3.21 – Backbone polar maps and global minimum conformation for 11-epi neopeltolide 3.29.

The 11-epi analogue primarily shows conformational differences from C8-C12. The 8-9-10-11 map now shows a loss of conformations with torsional values of 300°, while the 9-10-11-12 map loses its ones at 70-80°. Finally, the 10-11-12-13 polar map has the most significant changes, with all conformations from 180° to 280° vanishing completely while the formerly high energy torsional angles of 60-90° become preferred. These latter two polar maps, 9-10-11-12 and 10-11-12-13, are effectively mirror images of the natural ones, which would be congruent with the epimerization of the C11 stereocenter. Interestingly, the global minimum structure appears to orient the C11 methoxy in pseudo-axial fashion to avoid high-energy syn-pentane interactions (Figure 3.21).
The 11,13-\textit{epi} analogue has the most drastic conformational changes of these macrolides, extending from C9 to the macrocyclic ester oxygen. Its 9-10-11-12 map is effectively limited to a single available torsional angle, down from the three in the natural macrolide, and the 10-11-12-13, 11-12-13-O and 12-13-O-1 polar maps are exact mirror images of those from the parent compound. These changes to conformational preferences are appropriately manifested in the global minimum structure, which strongly resembles the wire structure proposed for Wright’s initial assignment of neopeltolide (Figure 3.22).
The 8,9-dihydroxy analogue presented an interesting change from the other analogues we investigated, representing an instance of addition to the macrocycle rather than epimerization. Conformational analysis of this analogue found that most of the polar maps remained relatively unchanged, with the 7-8-9-10 map seeing a narrowing of the preferred torsional angle. Major differences were seen from C9 to C12, with the low-energy torsional angles vanishing entirely from the 8-9-10-11 and 9-10-11-12 maps and the 10-11-12-13 map switching the low energy angle from 200° to 300°. Observation of the low energy conformer finds that it adopts a conformation reminiscent of the alternative proposed neopeltolide conformation, with the two hydroxy groups forming a hydrogen bonding network in concert with the C11 methoxy group (Figure 3.23). The introduction of two hydrophilic moieties to this analogue may also be playing a role in its
loss of biological activity, as this could be interfering with the molecule’s pharmacodynamics in addition to any conformational changes.

3.4 DISCON Analysis

Though the polar map analysis of neopeltolide’s macrolide suggests a relatively rigid molecule, it still demonstrates enough flexibility to require additional refinement with experimental data. To expedite the process of deconvoluting time-averaged NMR data, we chose to employ the previously-elaborated DISCON software developed by the Smith lab and discussed in Section 2.6.4.143

3.4.1 NMR Experiments

In order to properly utilize DISCON, we required a full panel of 1D and 2D NMR data of the neopeltolide macrolide. Fortunately, our computational experiments indicated that neopeltolide’s conformation should be similar in both aqueous and organic solvents, so we were able to make use of the complete set of high-field (600 MHz) \(^1\)H coupling constants collected in CDCl\(_3\) during the course of our lab’s formal synthesis.233 To complement these, we further collected 2D NMR data by running a ROESY experiment in CDCl\(_3\), which yielded a total of 36 cross-peaks. Using the C2 protons, these peaks were normalized and integrated to yield quantitative peak volumes.
Interestingly, the use of CDCl₃ as the NMR solvent allowed us to solve one of the interesting mysteries relating to the initial misassignment of neopeltolide’s absolute configuration. As mentioned previously, the use of d₄-methanol by Wright led to the H9 peak closely overlapping with one of the H8 peaks due to their virtually identical chemical shifts of 1.38 and 1.36 ppm, respectively. The switch to chloroform resulted in the H9 proton peak shifting to 1.5 ppm while the C8 peak stayed relatively similar at 1.39 ppm, which was distinct enough to distinguish between cross-peaks. As a result, the ROESY spectrum showed a distinct H8proS-H11 cross-peak but lacked an H9-H11 one, leading us to conclude that the cross-peak identified by Wright was a case of mistaken identity rather than evidence of an alternate conformer arising from C9-C12 flexibility (Figure 3.24). Curiously, Scheidt’s published NOESY spectrum does not appear to
possess an observable H9-H11 or H8-H11 peak (though this could be attributed to the relatively poor quality of their spectrum), but the accompanying work still reiterates the flexibility argument.\textsuperscript{253}

3.4.2 Identification of Neopeltolide’s Solution Conformation

With a complete set of NMR data in hand, we turned our attention towards identifying neopeltolide’s solution conformation with DISCON. Of the 36 NOE cross-peaks, we selected 9 for use with DISCON, along with 12 $^{3}$J$_{H-H}$ coupling constants. From our conformational searches of neopeltolide, we identified 20 unique structures to be used as the DISCON structure library.

From the early test runs, it quickly became apparent that neopeltolide does not exist as a mixture of solution conformations. Attempts to find the optimum number of clusters were relatively unsuccessful as total error was relatively consistent at each level. More troublesome was no ensemble solution at any of the explored clustering levels possessed a better RMS fit to the NMR data than the global minimum structure or the three other structures that differed only in the orientation of the ester relative to the macrolide (which cannot be evaluated via NMR). At cluster level 8, which possessed the minimum level of error, the global minimum-containing cluster represented over 90% of the solution contribution, with no other structures representing greater than 5%. Increasing the cluster level further resulted in this major cluster splitting into separate ones, the sum of which maintained $>$90% of the solution fit.
Figure 3.25 – Selected NMR variables used in DISCON and the two expected major conformers. Conformer B was not significantly represented in the data.

Closer examination of the NMR data, particularly the $^3J_{H-H}$ coupling constants, supports the theory of neopeltolide possessing a single dominant solution structure, Conformer A. The majority of these values are strongly weighted towards textbook values for gauche (2-5 Hz) and anti (8-11 Hz) orientations, which line up well with the observed orientations of protons on the global minimum structure. Additionally, several NOE cross-peaks that would indicate the second conformation posed in our preliminary work are notably absent, with the lack of an H8-H10proS peak being the most damaging
to its existence. The H11-H13 cross-peak is the only notable one that would lend credence to Conformer B, but the relatively close orientation of these protons in Conformer A would likely yield this medium-strength peak assuming minimal conformational changes in the structure over time. Visually, Conformer A also represents the only solution structure that completely avoids energetically-unfavorable syn-pentane interactions between the C9, C11 and C13 substituents in any combination, while every other conformer possesses at least one due to the constraints of the 14-membered ring (Figure 3.25).

3.5 Evaluation of Potential Neopeltolide Binding Sites on Cytochrome bc\textsubscript{1}

Following the elucidation of neopeltolide’s solution conformation, our attention turned to its proposed protein target, cytochrome bc\textsubscript{1}.\textsuperscript{224} The overall SAR profile painted by the various analogue studies and our conformational analysis work strongly suggest that neopeltolide’s macrolide conformation plays a role in its low-nanomolar cytotoxicity, even though its biological activity is highly dependent on the oxazole-containing side chain. As a result, its protein binding site should possess a noticeable pocket that would complement both the macrolide’s specific conformational preferences and the side chain’s necessary moieties, specifically its oxazole ring and carbamate tail. Since the binding site of neopeltolide is unknown, we believed that we could use the results of our conformational studies and the panel of known analogues to make a reasonable guess. With a known binding site, further insight could be gained into neopeltolide’s molecular mode of action, as well as helping with the design of future analogues.
3.5.1 Cytochrome $bc_1$

Cellular respiration depends on the function of a series of mitochondrial membrane-embedded protein complexes that make up the electron transport chain, which uses electron flow between its members to create an electrochemical proton gradient that serves as the motive force for the synthesis of ATP via oxidative phosphorylation.\textsuperscript{262} Cytochrome $bc_1$ is the third of these complexes, located in the inner mitochondrial membrane in eukaryotic cells. Structurally, this protein is a homo-dimeric complex that contains 11 units per monomer.\textsuperscript{263} Despite slight variations in composition depending on the source organism, only three subunits are essential for electron transfer: cytochrome $b$, cytochrome $c_1$, and the Rieske iron-sulfur protein. The cytochrome $b$ subunit contains two hemes ($b_L$ and $b_H$), cytochrome $c_1$ has one heme, and the Rieske subunit has a 2Fe-2S cluster. The two known active sites are located within cytochrome $b$, which is completely membrane-embedded, and are termed $Q_o$ (ubiquinol oxidation site) and $Q_i$ (ubiquinone reduction site).\textsuperscript{264}

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{Q_cycle.png}
\caption{Cartoon representation of the Q-cycle. From Vennam et al, Chem. Bio. Chem. 2013, 14, 1745-1753.}
\end{figure}
The function of cytochrome $bc_1$ is to catalyze the reduction of cytochrome $c_1$ via oxidation of coenzyme Q10, which can take the form of either ubiquinone or ubiquinol depending on its current oxidation state. In this process, called the Q cycle, two units of ubiquinol (QH$_2$) are oxidized at the $Q_o$ site of complex III, releasing four protons into the intermembrane space.$^{265-267}$ Two of the electrons released by this oxidation are transferred to the Rieske iron-sulfur protein and subsequently used to reduce an externally bound oxidized cytochrome $c_1$. The other two electrons are transferred through the cytochrome $b$ hemes and used to absorb two protons from the membrane space to reduce a unit of ubiquinone (Q) bound to the $Q_i$ site. The result is a net change of two protons across the membrane and the transference of reduced cytochrome $c_1$ to be used in complex IV (cytochrome $c$ oxidase), which drives both the proton gradient and the electron transport chain.

The crucial function of cytochrome $bc_1$ in cellular energy metabolism has made it a prime target for a wide variety of antibiotics and antifungals, with more than 20 synthetic compounds currently in widespread use.$^{268}$ Interestingly, known inhibitors of cytochrome $bc_1$ work in surprisingly mundane fashion despite the considerable size of the complex, disrupting the Q cycle by occupying either the $Q_o$ or $Q_i$ sites on the cytochrome $b$ subunit in place of coenzyme Q10. This has led to issues with resistance to many inhibitors due to target site mutations, making the development of agents that target areas outside these sites a tantalizing proposition.$^{269,270}$
Figure 3.27 – Biological data from Kozmin et al\textsuperscript{224} representing the inhibitory effects on cytochrome c oxidation (top left) and the difference absorption spectra.

Neopeltolide (and by association, leucascandrolide A) provides an interesting contrast to many of the known cytochrome \textit{bc}_{1} inhibitors. As shown earlier in Figure 3.4, most of these compounds are structurally similar to coenzyme Q10, which facilitates binding. However, virtually none of these compounds possess a macrolide as large as neopeltolide’s or structurally resemble it to a significant degree, opening the possibility that the polyketide possesses an alternate binding site or mode of action. Kozmin and co-workers carried out a series of spectroscopic studies that monitored changes in the oxidation states of cytochrome \textit{bc}_{1}, providing difference absorption spectra of yeast sonicated mitochondria in the presence of various inhibitors.\textsuperscript{224} Treatment with analogue
3.3 resulted in levels of reduced cytochrome \( b \) similar to the ones observed with other \( Q_0 \) inhibitors, even appearing to work in tandem with \( Q_1 \) inhibitor antimycin similar to the stigmatellin/antimycin mixture.

3.5.2 SiteMap Analysis of Cytochrome \( \text{bc}_1 \)

Our initial analysis focused on identifying all of the potential binding sites on cytochrome \( \text{bc}_1 \). In their attempt to discern the binding site of peloroside A, Schriemer and co-workers used the TR-NOESY conformation along with AutoDock to “blind” dock to a constructed \( \alpha-\beta-\alpha \) tubulin protofilament. In this exercise, the entire exterior surface of the protein was interrogated with the ligand conformation for potential binding sites based on docking energies. Their run revealed three sites of comparable docking energies, one of which matched the tubulin region found by the hydrogen-deuterium exchange study also reported in their work. Though no similar data exists for our own docking study, the limited number of known binding sites on cytochrome \( \text{bc}_1 \) led us to propose that we could use a similar docking method to evaluate potential binding sites for neopeltolide.

To keep all of our computational work within the same program, we decided to employ the Glide docking software found within the Schrödinger Suite. Unfortunately, Glide does not offer the same blind docking functionality that AutoDock does, but instead relies upon Schrödinger’s SiteMap routine. SiteMap defines a binding site as a set of points on a grid that are either contiguous or bridged by short gaps in solvent-exposed regions. To do this, the program places a 1Å grid of possible points around the entire protein, gauging the distance between each point and the van der Waals radius of the protein atoms to determine whether it is inside or outside the protein. Points
that are found to be outside the protein yet within good van der Waals contact of the receptor and sufficiently enclosed are gauged to be candidate points. If enough candidate points are within close range of one another and the group is solvent-exposed, the program deems the region they occupy as a potential binding site. The surface of each site is subsequently evaluated for hydrophobic and hydrophilic areas, with the latter being further subdivided into H-bond donor, H-bond acceptor, and metal binding regions. These sites are provided to the user in graphical form, and can be used as starting points to define docking grids for Glide.

Our first task was choosing a crystal structure for SiteMap analysis. Cytochrome $bc_1$ has been crystallized surprisingly often for a membrane protein and from several different biological sources, with the first structure reported in 1997. $^{263}$ We eventually settled on structure 3H1I ($Gallus gallus$, 3.53Å resolution)$^{264}$ from the RCSB PDB database, which would guarantee that SiteMap would at least find both the $Q_o$ and $Q_i$ sites as it has both stigmatellin and antimycin as co-crystallized ligands. SiteMap analysis of this protein complex using the default settings revealed a total of 10 possible binding sites, which includes both the $Q_o$ and $Q_i$ sites. The most interesting of these new regions was located on cytochrome $b$ in between the two known sites, which we have tentatively dubbed “Site 2” in accordance with its initial SiteMap identification. “Site 2” scored particularly well in SiteMap’s internal system, but the most striking feature of this site is its shape, which strongly resembles a macrocycle and side chain with a hydrophilic area at its tip (Figure 3.28). Subsequent evaluation of several other cytochrome $bc_1$ complexes, including 1NTM (apo, $Bos Taurus$, 2.4Å),$^{275}$ 1ZRT ($Q_o$/stigmatellin,
Rhodobacter capsulatus, 3.5Å),\textsuperscript{276} 1SQP (Q\textsubscript{o}/myxothaizol, Bos taurus, 2.7Å),\textsuperscript{277} and 1L0L (Q\textsubscript{o}/famoxadone, Bos taurus, 2.35Å),\textsuperscript{278} also found this same prospective site.

With our SiteMap results in hand, we began evaluating the prospective sites for their relative fit to neopeltolide. Glide was used to prepare a docking grid for each site and the Virtual Screening Workflow was used to dock our prepared neopeltolide conformer library to every site simultaneously with the high-throughput virtual screening (HTVS) function, the top poses from which were further refined with the standard precision (SP) and extra precision (XP) modes. The final docked poses ranked by GScore (Glide’s internal docking score, which predicts the non-covalent interaction strength

Figure 3.28 – SiteMap results on cytochrome b subunit illustrating the two major binding sites, known inhibitors, and the location of the newly-discovered “Site 2”.
between the ligand and protein) across all ten sites were saved and returned. These poses showed that the $Q_0$ site possessed the highest scoring pose (-9.624) with “Site 2” not far behind (-8.907). Interestingly, the $Q_i$ site’s top ranked pose was considerably lower than the majority of the “Site 2” poses (-7.099), and the remaining seven sites failed to produce any poses ranked higher than -6. While hardly conclusive, this screening established both the $Q_0$ site and “Site 2” as good candidates for further docking analysis.

Figure 3.29 – Global SiteMap output for the 3H1I monomer (left) and the results of a “blind” docking to each of them (right).

3.5.3 Induced Fit Docking of Neopeltolide

To examine the proposed binding sites in greater detail, we needed to investigate the possible binding poses of neopeltolide. First, we switched our model system from the 3H1I crystal structure to the 1L0L one, which has famoxadone (FMX) bound to the $Q_0$ site rather than stigmatellin and possesses better resolution than the former. Famoxadone has an oxazolidine ring (see Figure 3.4), which we hypothesized could be analogous to
the oxazole ring within neopeltolide’s side chain. SiteMap analysis of this complex was also successful in locating “Site 2” for subsequent docking.

Molecular docking of our neopeltolide library to the \( Q_0 \) site and “Site 2” provided several interesting poses with exceptional GScores. In the \( Q_0 \) site the side chain of neopeltolide appears to act as a coenzyme Q10 mimic, burrowing into the binding pocket while the macrolide remains outside. In “Site 2”, the entirety of neopeltolide lies flush against the cleft in reasonably good fit with the SiteMap profile. However, while these poses serve as useful starting points they did not offer significant non-hydrophobic contact with the binding pocket, with the “Site 2” pose providing only two instances and the \( Q_0 \) pose providing none. While “Site 2” is largely hydrophobic due to its membrane exposure, the known necessity of the native side chain makes it highly likely that there should be several hydrophilic contacts with the carbamate and oxazole ring at the very least. Given the presence of such an area in the SiteMap search, it opened the possibility that the binding pocket needed to accommodate neopeltolide to reflect a bound complex.

Standard virtual docking studies hold the receptor rigid while allowing the ligand to move, which could potentially bias the receptor against a dissimilar ligand. To simulate the “induced fit” that a ligand and receptor would experience biologically and hopefully obtain a more accurate complex, we used the induced-fit docking (IFD) protocol within Schrödinger to dock neopeltolide to our proposed sites.\textsuperscript{279,280} This protocol docks a ligand to a receptor with Glide using a softened van der Waals radii scaling and follows it with a Prime side-chain prediction and minimization on residues within 5Å of the pose to reflect the induced fit. Subsequently, the same ligand is re-docked into the new receptor and scored using standard Glide settings.
Using the IFD protocol on “Site 2” with extended sampling yielded 35 poses. The majority of these poses reflect the best pose from the first docking run, with neopeltolide lying perpendicularly in the pocket. Additionally, the best IFD pose reflects some of what is known about neopeltolide’s SAR. The C13 $n$-propyl tail, found by Fuwa et al to be crucial for low-nanomolar inhibition,\textsuperscript{257} is facing into the hydrophobic pocket while the unnecessary C9 methyl faces outward. From the side chain perspective, the orientation of the alkyl chain between the macrolide and oxazole matches what would be expected of the equipotent $Z,E$-dienoate \textbf{3.12} (Figure 3.6).\textsuperscript{254} The oxazole ring appears to be participating in both $\pi-\pi$ stacking and hydrogen bonding, while the carbamate tail also participates in hydrogen bonding. Different poses disagree on the residues that participate in side chain interactions, with TYR103, TYR104 or TRP326 interacting with the oxazole ring and ARG40, GLN72, GLN322, TRP30, or SER28 interacting with the carbamate in some fashion.
Conversely, the binding at the Qo site was less encouraging. While the docking score was improved only a single hydrophilic contact was made at most in any of the poses, with little consistency in the participating ligand moiety or receptor residue. The best scoring pose did forge a hydrogen bonding contact between the oxazole ring and GLU271, which is analogous to the connection between famoxadone’s oxazolidine ring and the same residue, but was the only one of the 19 poses to do so. Other poses found π-π stacking interactions for the oxazole ring or a hydrogen bonding interaction for the carbamate, but no single pose illustrated the simultaneous need for both moieties that was demonstrated in neopeltolide’s SAR profile. Additionally, with the bulk of the molecule hanging out of the binding pocket in every pose, binding at the Qo site would not explain any of the macrolide SAR or the apparent conformational requirements for low nanomolar inhibition.
3.5.4 Molecular Dynamics of Neopeltolide-Bound Complexes

Although IFD provides for a better docking fit than a standard docking run, the number of different binding poses that scored similarly well despite different residue interactions does not leave a clear picture for a potential binding mode at either site. In the peloruside A study by Schriemer and co-workers,\textsuperscript{271} they followed up their blind docking studies by using the generated poses in molecular dynamics (MD) simulations. Subsequently, they determined the binding energies using MM-GBSA to gauge the potential binding for each site. As such, we decided to use the top poses generated from IFD as initial conformations for our own MD simulations to determine how well they fit into the complex.

To be able to use our structures directly from other Schrödinger programs, we decided to use D.E. Shaw Research’s Desmond\textsuperscript{281} for MD simulations. Our system was built using the 1L0L complex from the Orientations of Proteins in Membranes (OPM) database,\textsuperscript{282} which provides proteins that have been pre-positioned with respect to the hydrocarbon core of the lipid bilayer. The SiteMap and IFD protocols were repeated as before on one of the protein monomers to provide the initial neopeltolide-bound conformations, leaving famoxadone bound as normal to the Q\textsubscript{o} site in the other monomer. This was done to provide a visual contrast to potential conformational changes that neopeltolide might cause over the course of the simulation.

Next, the MD system was constructed using Desmond’s System Builder, using its built-in functions to install a POPC lipid bilayer around the complex. While the precise mechanism of cytochrome’s normal operation is still a matter of debate, we wanted to mimic the theoretical biological environment as closely as possible. A periodic boundary
box was constructed 10Å around the complex on all sides, consisting of the lipid bilayer in the middle and SPC waters otherwise. Chlorine ions were also placed throughout the system to neutralize charges. The system was equilibrated using Schrödinger’s membrane relaxation protocol, which gradually raises the temperature of the system to 300K before performing a series of short (0.2-0.6 ns) NPT simulations on the different parts of the system. Both the Q\textsubscript{o}/neopeltolide and “Site 2”/neopeltolide complexes were simulated for 40 ns at 300K using the OPLS-AA force field and NPT ensemble. The Langevin thermostat and baryostat were employed, along with a 2 fs timestep throughout the simulation.

Analysis of the total kinetic and potential energies of the “Site 2”/neopeltolide simulation found that it levelled out approximately halfway through the run, indicating that the system achieved equilibrium. The complex saw some initial movement of both the protein and ligand according to a plot of root mean square deviation (RMSD) of the backbone carbons over time, but appeared to stabilize at around 20 ns. Visual analysis of neopeltolide finds that the macrolide demonstrated virtually no movement while the side chain experienced minor fluctuations. The oxazole ring retained its π-π stacking interaction with TYR103 while maintaining a hydrogen bonding interaction with TRP30. The carbamate tail is a bit more promiscuous, with the carbonyl first forming a hydrogen bond with TYR104 but changing to TRP30 at brief intervals.
To gain a better idea of how well neopeltolide bound to “Site 2”, we used Schrödinger’s Prime module to perform MM-GBSA calculations on the bound complex. This method calculates the free energy difference between the bound and unbound state of two solvated molecules, deriving the binding free energy $\Delta G_{\text{bind}}$ by breaking down the thermodynamic cycle of binding into component terms that can be calculated with the same force field energies and GB/SA solvation methods used in our conformational searches.\textsuperscript{283} Prime MM-GBSA found that neopeltolide binds strongly to “Site 2”, with an average $\Delta G_{\text{bind}}$ of -77.1 over the course of the simulation; for comparison, famoxadone had an average $\Delta G_{\text{bind}}$ of -60.5 over the same span.
The Qₐ/neopeltolide complex, on the other hand, was markedly different. While the total kinetic and potential energy levelled out at approximately the same time as the “Site 2” complex, RMSD analysis found significant perturbation in both the ligand and protein complex. Looking at the trajectory, neopeltolide appears to move wildly within the pocket for the first 8 ns before extending itself through the site entirely. Over the remainder of the simulation, the ligand burrows into the protein, causing significant displacement of the Rieske iron-sulfur protein and forcing it to virtually dissociate from the remainder of the complex prior to achieving equilibrium. The instability of the bound complex suggests that neopeltolide is not a particularly good fit for the Qₐ site within the current understanding of cytochrome bc₁ inhibition.

Figure 3.33 – Snapshots from Qₐ/neopeltolide MD simulation illustrating Rieske iron-sulfur protein deviation on the neopeltolide-bound monomer (green) vs. the famoxadone-bound one (red). Also shown is the RMSD experienced by the system (blue = protein, red = neopeltolide).
3.5.5 “Site 2” Docking of Analogues

Our goal at the outset of this project was to dock the existing panel of neopeltolide analogues developed by Fuwa and co-workers\textsuperscript{255,257} into any proposed binding sites as a measure of their validity. Since they used the same A549 assay to give their biological results for each analogue, we hoped that a relative comparison could be drawn between the natural logs of analogue IC\textsubscript{50} values and MM-GBSA binding energies through the relationship of Gibbs free energy to binding constants ($\Delta G = -RT \ln K_{eq}$). Conformational searches were performed on each of Fuwa’s macrolide and side chain analogues to provide a general library for docking, which was docked into the receptor derived from the “Site 2”/neopeltolide MD simulation using Glide. The resulting top scoring poses were evaluated using Prime MM-GBSA calculations as before.

Figure 3.34 – Graphs comparing Prime-MMGBSA binding energies to the natural log of analogue IC\textsubscript{50} values. The “All Analogues” chart also includes the Fuwa side chain analogues illustrated in Figure 3.8.
Unfortunately, comparison of MM-GBSA binding affinities to reported IC<sub>50</sub> values did not yield a reliable correlation. As shown in Figure 3.34, comparing the entire panel of analogues yielded an R<sup>2</sup> fit of only 0.0701. We though that the flexibility observed in the side chain during the MD simulation could potentially be a factor, as MM-GBSA does not account well for conformational entropy, so we next limited the panel to the macrolide analogues only. This improved the fit somewhat (R<sup>2</sup> = 0.1099), but not to the degree where we could assert a direct correlation.

Further issues were also raised by the results of the neopeltolide docking. Several analogues were found to bind with nearly identical GScores when the macrolide is reversed within the pocket (C9 methyl and C11 methoxy pointing inwards, C13 n-propyl chain exposed to solvent), which either calls into question our initial binding pose or raises the possibility of multiple binding modes.

3.6 Total Synthesis of Neopeltolide and 2-Methyl Neopeltolide

3.6.1 Taylor Formal Synthesis

Our lab’s formal synthesis of neopeltolide was completed by Rendy Kartika in 2008, using the macrolide synthesized by Panek as the final target. This approach employed a novel intramolecular ether transfer reaction to establish the stereochemistry of the tetrahydropyran moiety and a radical cyclization to complete its construction.
Scheme 3.1 – Conditions for Taylor’s synthesis of β-alkoxyacylate 3.54 en route to the macrolide core of neopeltolide.

a) BH$_3$·SMe$_2$; b) (OMe)MeNH$_2$Cl, iPrMgCl; c) DMP, 86% (over 3 steps); d) 3.47, BF$_3$·OEt$_2$; e) BOMCl, Hünig’s base, 55% (over 2 steps); f) SmI$_2$, PhCHO (dr > 20:1); g) Me$_3$OBF$_4$, Proton Sponge, 73% (over 3 steps); h) ICl; then Na$_2$S$_2$O$_3$, 71% (dr > 20:1); i) ethyl propiolate, PBu$_3$, 98%.

Similar to Panek’s approach, the synthesis began with methyl (R)-(+)3-methylglutarate 3.45, which was transformed to the Weinreb amide 3.46 over three steps. Asymmetric allylation at the more electrophilic aldehyde moiety with Soderquist’s chiral bicyclocdecane-allylborane reagent 3.47 introduced the homoallylic alcohol, which in turn was BOM-protected to give ether 3.48. The C12-C16 segment of neopeltolide began with β-hydroxysulfide 3.49, which was transformed to the dilithio species 3.50 and
coupled to the Weinreb amide using Rychnovsky’s method\textsuperscript{285} to yield the β-hydroxy ketone 3.51. The C11 stereocenter and transformation of the C13 hydroxyl to a benzoate ester were realized simultaneously with an Evans-Tischenko reaction,\textsuperscript{286} and the resulting C11 hydroxyl was methylated to yield methyl ether 3.52.

\[
\begin{align*}
\text{3.54} & \xrightarrow{a} \text{3.55} \\
\text{3.55} & \xrightarrow{b} \text{3.56} \\
\text{3.56} & \xleftarrow{c, d} \text{3.57}
\end{align*}
\]

Scheme 3.2 – Conditions for completion of the neopeltolide macroclide 3.57. a) AIBN, \(n\text{Bu}_3\text{SnH}\); 95\% (dr 19:1); b) KOH, MeOH, 95\%; c) TCBCl, Et\(_3\)N, DMAP, 87\%; d) H\(_2\), Pd/C, 100\%.

Next, ether transfer methodology was used to directly install the C5 stereocenter as a benzyl-protected ether via treatment with iodine monochloride.\textsuperscript{287} Following aqueous workup, the 1,3-\textit{syn}-diol monoether 3.53 was liberated in good yield with excellent stereocontrol. The remainder of neopeltolide’s carbon skeleton was achieved with the tributylphosphine-promoted conjugate addition of the alcohol to ethyl propiolate to yield 3.54. The tetrahydropyran core was formed through radical cyclization with AIBN and \(n\)-
Bu₃SnH in refluxing toluene to give 3.55, followed by exposure to KOH and methanol to simultaneously saponify the ethyl ester and cleave the C13 benzoate ester. Finally, Yamaguchi macrolactonization of the seco acid 3.56 followed by hydrogenolysis with Pd/C catalyst to remove the C5 benzyl ether completed the macrocyclic core of (+)-neopeltolide. Overall, the synthesis required 14 steps in the longest linear sequence, with an overall yield of 18.9% starting from 3.45. This route also proved to be extremely scalable, ultimately affording over 100 mg of macrolide.

3.6.2 Synthesis of 2-Methyl Macrolide

With our polar map analysis strongly suggesting that C2 substitution of the neopeltolide macrolide should have little to no effect on biological activity, we decided to test this theory by synthesizing 2-methyl neopeltolide for biological studies. Fortuitously, as C2 protons are directly adjacent to the ester carbonyl we envisioned that a simple enolate addition could be achieved with the completed macrolide. After protecting the C5 hydroxyl with TBSOTf and 2,6-lutidine to yield 3.58 in 99% yield, conditions for addition were screened. Eventually, 5 equivalents of NaHMDS and 6 equivalents of methyl iodide bore a successful reaction, giving the TBS-protected analogue in 82% yield. NMR experiments showed the disappearance of both C2 protons and the introduction of a doublet of quartets in the same region, with the doublet portion possessing a coupling constant consistent with an axial-axial orientation to the C3 proton (9.8 Hz). This indicates that the methyl group was installed equatorial to the macrolide ring, which would be the (R)-2-methyl product. A COSY spectrum of the product also finds that this peak correlates to a new 3H doublet at δ 1.08 ppm, which would be the methyl in question.
Scheme 3.3 – Synthesis of (R)-2-methyl macrolide 3.58.

Interestingly, this reaction gives only a single product with an excess of 20:1 dr. The reason for this likely stems from the fact that only the Z-enolate can be formed with the macrolide, as the E-enolate would be an extremely high-energy intermediate due to ring strain (Figure 3.35). Since the macrocycle represents a significant amount of steric bulk the electrophilic methyl iodide should preferentially approach from outside the ring. Floreancig and co-workers used this bulk to their advantage in their synthesis of neopeltolide, where the final hydrogenation to establish the C9 stereochemistry proceeded with high facial control due to the conformation of their precursor alkene.234

Figure 3.35 – Comparison of possible C2 enolates. Major conformer (left), Z-enolate (middle), E-enolate (right).
Similarly, our addition to C2 exclusively yielded the (R)-2-methyl product, as a result of the exterior face of the macrocycle being more sterically accessible (Figure 3.36). Interestingly, attempts to racemize this stereocenter by exposure to similar conditions were unsuccessful, as were attempts to either dimethylate C2 or convert the remaining C2 proton to a deuterium. This strongly suggests that the remaining C2 proton is far less susceptible to basic conditions, which could either be a result of steric hindrance with the relatively bulky NaHMDS base or a change in pKa due to additional alkyl substitution.

![Figure 3.36 – Possible explanation for the observed stereoselectivity of the C2 macrolide addition (below). Floreancig’s hydrogenation of 8,9-dehydro macrolide (above) is shown to illustrate a similar sterically-driven facial preference.](image)

3.6.3 Side Chain Synthesis and Completion

With the different macrolides in hand, our attention turned to the oxazole-containing side chain to complete the total synthesis of neopeltolide and (R)-2-methyl
neopeltolide. At the outset of the project, our goal was to develop a modular route that would allow us to test different aromatic functionalities and tails, as these analogues had yet to be published. However, while attempting to find a way to couple these functionalities we were undercut by the independent efforts of Floreancig\textsuperscript{218} and Fuwa,\textsuperscript{255} who established the necessity of the native side chain as a prerequisite for potent biological activity. As our own route proved to be unfruitful, we instead decided to make the side chain with the route developed by Leighton and co-workers in their total synthesis of leucascandrolide A.\textsuperscript{243}

Scheme 3.4 – Route to side chain 3.65 devised by Leighton and co-workers.\textsuperscript{243}

The side chain began with the synthesis of carbamate 3.59 from propargylamine and methyl chloroformate. Deprotonation of this substrate with nBuLi followed by
bubbling CO₂ gas through the mixture afforded the ynoic acid, which was promptly subjected to Lindlar reduction to give the Z-enoic acid 3.60. Coupling to L-serine methyl ester via the mixed anhydride formed with isobutyl chloroformate yielded the amide 3.61. This amide was converted to the oxazole 3.62 by employing a one-pot method using diethylaminosulfurtrifluoride (DAST), DBU and BrCCl₃. The ester was reduced using DIBAL-H to give the alcohol, which was subsequently treated with CBr₄ and PPh₃ to give the bromide 3.63. Stille coupling with vinyltributyltin gave the allyloxazole 3.64, and hydroboration of the alkene with 9-BBN followed by Swern oxidation of the resultant alcohol gave the final side chain aldehyde 3.65. Overall, our synthesis yielded approximately 40 mg of product.

Scheme 3.5 – Completion of the synthesis of neopeltolide 3.1b and (R)-2-methyl neopeltolide 3.68.
Coupling of the side chain to the neopeltolide macrolide 3.57 and the 2-methyl analogue 3.58 was accomplished using a Still-Gennari olefination as reported by Panek and co-workers. The macrolide alcohols were acylated with bis(2,2,2-trifluoroethyl)phosphonoacetic acid to give the phosphonoacetates 3.66 and 3.67, which were immediately deprotonated with KHMDS at -78°C in the presence of 18-crown-6 ether. Treatment of the resultant anions with 3.65 at -85°C gave the final products neopeltolide 3.1b and 2-methyl neopeltolide 3.68. Spectroscopic data of our synthetic neopeltolide matched the set described by Panek and co-workers, and ¹H data for 3.68 established the formation of the olefin in approximately a 4:1 Z:E ratio roughly consistent with the Still-Gennari olefination used to form neopeltolide. Mass spectrometry for 3.68 also matched the expected value for 2-methyl neopeltolide.
4.1 Purpose

The purpose of this chapter is to describe the marine-derived polyketide natural products zampanolide and dactylolide and our efforts towards identifying their bioactive conformation. The isolation, biological activity and relationship between the two macromolecules will be covered, followed by a brief overview of synthetic work towards their construction and a critical evaluation of previous conformational studies. High-field NMR experiments and molecular modeling work indicated that zampanolide and dactylolide’s shared macrolide core prefers to adopt three major conformational families in solution. The conformational preferences exhibited by these families will be presented, along with a comparison to the tubulin-bound conformation of zampanolide and an outline of our plans for conformationally-similar analogues of zampanolide.

4.2 Background

4.2.1 Zampanolide Isolation and Structure

In 1996, Tanaka and co-workers reported the discovery of a novel natural product, termed zampanolide, from the marine sponge *Fasciospongia rimosa* off the coast of Okinawa, Japan. The isolation involved extracting a 480 g sample of wet sponge with acetone followed by ethyl acetate, which gave 2.87 g of crude material. Purification of
this material via silica-gel chromatography yielded over a gram of the natural product latrunculin A and a minute 3.9 mg of zampanolide. A decade later, Field and co-workers isolated zampanolide again, this time from a sample of *Cacospongia mycophijiensis* collected off the coast of ‘Eua, Tonga. As with Tanaka’s isolation, their process yielded several natural products in addition to 1.6 mg of pure zampanolide from 341 g of frozen sponge.

After their initial isolation, Tanaka and co-workers began determining the structure of their newly-discovered natural product through a combination of mass spectrometry and multiple 1D and 2D NMR experiments. Their data revealed that the core of zampanolide consists of a 20-membered macrolide containing a cis-2,6-tetrahydropyran, three (E)- and one (Z)-olefinic bonds and three stereocenters. Additionally, zampanolide possesses an unusual N-acyl hemiaminal side chain at C19, a structural feature that has been found in a limited number of bioactive natural products. This side chain contains zampanolide’s fourth and final stereocenter at C20. While Tanaka *et al* were able to assign the relative configurations of the macrolide stereocenters via an NOE network between C11, C15, Me23 and C19, the absolute stereochemistry and the configuration of the C20 hemiaminal eluded them.

![Figure 4.1 – Zampanolide and dactylolide.](image-url)
4.2.2 Dactyloolide Isolation and Structure

In 2001, Riccio and co-workers isolated dactyloolide from the marine sponge *Dactylospongia* sp. off the coast of the Vanuatu islands in the South Pacific Ocean. Their raw sponge extract yielded sizeable amounts of mycothiazole, laulimalide, isolaulimalide and latrunculin A with medium pressure liquid chromatography. However, further purification of several of their impure fractions with reversed-phase HPLC afforded a small amount of dactyloolide as an amorphous solid.

Following isolation, Riccio and co-workers elucidated the structure of dactyloolide. The identification of dactyloolide’s molecular formula revealed an interesting problem during HRMS analysis. When the compound was ionized in methanolic solution, two major ion peaks were observed at $m/z = 417$ and $m/z = 439$, which were respectively assumed to be the proton and sodium adducts. However, additional ESIMS/MS experiments revealed that both these peaks readily lost a molecule of MeOH (32 a.m.u.) to give stable ions at $m/z = 385$ and $m/z = 407$. Repeating these experiments in deuterated methanol gave a prominent sodium adduct peak at $m/z = 443$ [M + CD$_3$OD + Na]$^+$ that lost a molecule of CD$_3$OD to afford a stable ion at $m/z = 407$, indicating that methanol was adding to the structure of dactyloolide (Figure 4.2).

![Figure 4.2 – Reaction of methanol with dactyloolide.](image)
Further analysis using methods similar to those of Tanaka et al. revealed that dactylolide consists of a macrolide core virtually identical to that of zampanolide, though Riccio and co-workers did not initially assign the C19 stereocenter. Dactylolide lacks the hemiaminal side chain of zampanolide, however, instead possessing an aldehyde at C19 that acts as the site of methanol addition observed in MS experiments. Additionally, natural dactylolide has been proposed to have the opposite absolute stereochemistry of zampanolide, possessing a positive $[\alpha]_D$ of +30° as opposed to zampanolide’s -165.8°.

4.2.3 Biological Activity and Mode of Action

In addition to its isolation and tentative structure, Tanaka and co-workers also reported that zampanolide displays potent ($IC_{50} = 1-5$ nM) cytotoxicity against a variety of cancer cell lines, including P388 (leukemia), HT29 (colon), A549 (lung) and MEL28 (melanoma). Material tested in 2009 by Field and co-workers expanded this cytotoxic activity to HL-60 and A19 cancer cells as well. Field et al. also reported the first studies on the mode of action of zampanolide at the same time, which had yet to be performed owing to the scarcity of material. Their work concluded that the compound exhibits antimitotic action, causing arrest in the G2/M phase of the cell cycle. Additionally, treatment with zampanolide showed significant amounts of microtubule bundling in interphase cells and multiple asters in dividing cells, suggesting that microtubules were its protein target. Finally, zampanolide caused a dose-dependent shift of soluble tubulin to its polymerized form, which is commonly seen in microtubule-stabilizing agents such as paclitaxel and discodermolide. Similar to several of the MSAs discussed in Chapter 1, zampanolide showed no sensitivity to human ovarian cancer cells that display resistance through overexpression of the P-glycoprotein pump.
For dactylolide, Riccio and co-workers reported that it displayed somewhat more modest cytotoxicity than zampanolide against L1210 (lymphatic leukemia) and SK-OV-3 (ovarian carcinoma) cancer cell lines, with 63% and 40% inhibition at 3.2 μg/mL respectively. The unnatural enantiomer, (–)-dactylolide, prepared by the Jennings laboratory in 2008 was tested against the 60-cell NCI panel, which found that it exhibited GI$_{50}$ values in the nanomolar range (25-99 nM) against HL-60, K-562, HCC-2998 and SF-539 cancer cell lines. In direct comparison with (–)-zampanolide against a handful of cell lines dactylolide was found to be 10- to 1000-fold less active, suggesting that the hemiaminal side chain plays a key role for biological activity. Field and co-workers later concluded that dactylolide also displays microtubule-stabilizing activity, arresting A549 cancer cells in the G2/M phase and producing long microtubule bundles.

Both zampanolide and dactylolide are of particular interest owing to their ability to bind tubulin and induce microtubule polymerization under conditions where it would otherwise not occur. Field and co-workers reported that zampanolide promoted microtubule assembly at a concentration of only 4.1 μM in a buffer designed to inhibit tubulin polymerization. Additionally, they noticed that they were completely unable to recover either compound from their microtubule pellets, suggesting that both polyketides bind irreversibly to microtubules. The only other natural products known to bind tubulin in covalent fashion are pironetin, cyclostreptin and taccalonolide, with only cyclostreptin suspected of binding to the taxane site. Field and co-workers later used competitive inhibition studies, Flutax-2 displacement experiments and targeted mass spectrometry experiments to conclude that zampanolide and dactylolide bind to the taxane site at residue H229, though a small amount also reacted with residue N228.
Interestingly, they also found that zampanolide and dactylolide react at different rates following non-covalent binding, with the former reacting immediately while the latter required nearly 4 h for complete conversion and appeared to be in fast exchange with the medium. This suggests that the hemiaminal side chain plays not only a role in activity but also in the kinetics of binding.

![Crystal structure of zampanolide bound to β-tubulin](image)

Figure 4.3 – Crystal structure of zampanolide bound to β-tubulin (RSC PDB: 4I4T). Hydrogen bonding interactions shown in green.

Field and co-workers also made an attempt to determine the binding pose of zampanolide within the taxane pocket, using AutoDock along with an electron crystallography-derived structure of α,β-tubulin (RSC PDB: 1JFF) to produce several potential configurations that would reflect covalent binding at potential electrophilic
moieties. Though they were unable to definitively identify the correct pose, they did establish that each of their binding modes brought several favorable contacts to the M-loop of tubulin. This led them to conclude that zampanolide and dactylolide stabilize microtubules in similar fashion to paclitaxel and the epothilones, which also modify the M-loop. The ultimate binding pose of zampanolide to tubulin was reported in 2013 by Prota and co-workers, who published a high-resolution (1.8Å) co-crystal structure of the bound complex. This structure demonstrated that the zampanolide core forms a covalent bond between H229 and its α,β-unsaturated enone, which was predicted by Field et al. It also found that the hemiaminal side chain of zampanolide makes two crucial connections with T274, a residue on the M-loop that is also involved with epothilone’s microtubule-stabilizing mechanism (Figure 4.3).

4.2.4 Biological Relationship between Zampanolide and Dactylolide

With the obvious structural similarities between the two polyketides, the community has naturally wondered about their biosynthetic relationship. Although dactylolide and zampanolide share the same macrolide core, the opposite absolute stereochemistry of the two implies an alternate biosynthetic pathway. However, doubts have been cast as to the accuracy of the natural compound’s optical rotation; several research groups have prepared both the natural and unnatural enantiomers of dactylolide and reported specific rotation values between 134° and 258°, values significantly greater than 30°. With this discrepancy, it raises the possibility that the initial value was recorded incorrectly and the absolute stereochemistry of natural dactylolide is the same as that of (−)-zampanolide.
As dactylolide merely lacks the hemiaminal side chain of zampanolide, several groups have postulated that it might be a degradation product of the latter. In 2002, the Smith group demonstrated that (+)-dactylolide could be obtained through thermolysis of (+)-zampanolide at 85°C. However, as the reaction proceeds through a pseudo retro-ene process, the necessary heat would be unlikely to be present in a marine sponge’s natural environment. Treatment of zampanolide with strong base was also unsuccessful in yielding dactylolide. Tanaka and co-workers later tried subjecting zampanolide to the isolation conditions used by Riccio laboratory, but were unable to find dactylolide by LC-MS. They also were unable to find any trace of dactylolide in a fresh sample of Fasciospongia rimosa, suggesting that dactylolide is a naturally occurring entity in its own right. While our laboratory has proposed a biosynthetic pathway for the formation of zampanolide, the exact pathways that give rise to both it and dactylolide, and their relationship to one another, are still under investigation.

4.2.5 Total Synthetic Efforts

The potent biological activities displayed by dactylolide and zampanolide and the relatively few number of stereocenters they possess have made them attractive targets for the synthetic community. To date, syntheses of both natural and unnatural dactylolide have been reported by eleven research groups. Zampanolide, conversely, has had considerably fewer reports of successful synthesis despite being the more active of the two compounds, which is likely due to the difficulty in installing and maintaining the C20 stereocenter. The first synthetic route published by the Smith laboratory in 2002 found that this stereocenter epimerized upon final deprotection of the alcohol, and attempts by other groups to install the hemiaminal side chain to the dactylolide core has
led to relatively low yields and diastereoselectivity. The synthetic efforts towards dactylolide and zampanolide have been covered in reviews elsewhere.\textsuperscript{303}

4.2.5.1 Taylor Lab Synthesis of Dactylolide and Zampanolide

Scheme 4.1 – Taylor retrosynthesis of zampanolide.

Our own lab’s synthesis, spearheaded by Matt Wilson, of dactylolide and zampanolide was based on a convergent strategy starting from (R)-aspartic acid, which targeted the formal intermediate 4.4 achieved by the Jennings laboratory in their total synthesis of (−)-dactylolide.\textsuperscript{292} This route is highlighted by an organocatalytic amidation, ring opening of a vinyl epoxide, a one-pot cross-metathesis/Horner-Wadsworth-Emmons
olefination, and our lab’s ether transfer methodology. Full details of this route, its
evolution, and contrast to other total syntheses are covered elsewhere.\textsuperscript{301}

\begin{align*}
&\text{(R)-aspartic acid} \\
\xrightarrow{\text{PBu}_3, \text{DCM, 97\%}} &\text{MeO} \quad \xrightarrow{\text{Et}_3\text{B, EPHP}} &\text{MeO}_2\text{C} \\
\end{align*}

5 steps, 56\% \quad 1\text{Cl}, -78^\circ\text{C} \quad \text{ICl, -78^\circ\text{C}} \quad \text{PhMe, H}_2\text{O/DIPA} \\
\text{I} \quad \text{-CH}_2\text{O} \quad 58\% \text{ yield, 8:1 dr} \\

\begin{align*}
\text{OPBB} &\quad \text{OPBB} \\
\text{BnO} &\quad \text{BnO} \\
\text{4.6} &\quad \text{4.7} \\
\end{align*}

\begin{align*}
\xrightarrow{\text{EIOH, 88\%}} &\text{MeO}_2\text{C} \\
\text{OBn} &\quad \text{OBn} \\
\text{4.8} &\quad \text{4.9} \\
\end{align*}

\begin{align*}
\xrightarrow{\text{3 steps, 76\%}} &\text{OBn} \\
\text{4.5} &\quad \text{OBn} \\
\end{align*}

\text{Scheme 4.2 – Synthetic route to pyran 4.5 for cross-coupling}

Briefly, BOM ether 4.6 was first synthesized over 5 steps from (R)-aspartic acid.
The 1,3-\textit{syn} diol monoether motif embedded within the \textit{cis}-2,6-trisubstituted pyran
precursor 4.5 was constructed via our electrophile-induced ether transfer methodology,
using iodine monochloride on an 8-gram scale to give 4.7. Next, this compound was
converted to the \(\beta\)-alkoxyacrylate 4.8 so that the vinyl iodide could be used with a radical
cyclization to yield the pyran methyl ester 4.9, similar to what was used in our lab’s
synthesis of neopeltolide.\textsuperscript{233} The methyl ester was subsequently converted to the terminal
olefin over three steps.
Scheme 4.3 – Route to (−)-dactyloolide 4.14 and (−)-zampanolide

The C16-C20 fragment and E-trisubstituted olefin was built using a tandem cross-metathesis/HWE olefination to yield dienoate 4.10 as a single geometrical isomer. Subsequent reduction of the ester followed by Sharpless epoxidation gave the vinyl epoxide 4.11, which was selectively opened and acetalized to yield 4.12. The C9 and C13 olefins were next installed over three steps through removal of the benzyl ethers, bis-oxidation and bis-Wittig olefinations, and the targeted intermediate 4.4 was finally obtained through hydrolysis of the acetonide and selective ether protection. Final synthesis of (−)-dactylolide 4.14 was achieved through use of the work of Jennings et al over four steps, coupling their intermediate 4.3 via Yamaguchi macrolactonization and
ring-closing metathesis followed by oxidation with DMP to yield the C20 aldehyde and the C7 enone. The conversion of dactylolide to zampanolide was achieved through the use of Ghosh’s protocol,\textsuperscript{304} which unfortunately highlighted the reason for limited synthetic zampanolide amounts with its relatively low yield and 3:1 ratio of zampanolide and its C20 epimer. Overall, our route provided sufficient material for use in biological studies and the conformational analysis outlined in this chapter.

4.3 Previous Conformational Work

4.3.1 Porco: Conformation of the Hemiaminal Side Chain

In 2002, the Porco laboratory reported on their efforts towards the total synthesis of zampanolide, namely their preliminary work on the synthesis of the hemiaminal side chain.\textsuperscript{305} While they were working on the hydrolysis of their model compounds to the hemiaminal, they saw a noticeable difference in stability depending on the presence of a sorbic side chain (Scheme 4.4). The compounds such as 4.15 containing this functionality were significantly more stable than those that did not (4.16), the latter of which would readily decompose or react unfavorably under solvolysis conditions. Their hypothesis was that this difference was due to the hydrogen bonding between the hemiaminal nitrogen and the ester carbonyl in 4.15, which would discourage the formation of the iminium ion species leading to decomposition.
Scheme 4.4 – Hemiaminal model substrates prepared by the Porco laboratory

To test this theory, they took $^1$H NMR spectra of a diastereomeric mixture of their model compound 4.15 in acid-free CDCl$_3$ after exhaustive D$_2$O exchange. From this, they were able to deduce that the epimeric H$_a$ protons had coupling constant values of 2 and 7 Hz, respectively. Modeling both of these compounds utilizing an MM2 force field provided a set of low energy conformations for each diastereomer that would reflect their proposed hydrogen bonding network (Scheme 4.5). Interestingly, the predicted coupling constant values for each diastereomer closely matched the ones they observed experimentally, strongly implying that they adopt the predicted conformations.
Scheme 4.5 – Representative hydrogen bonding networks in hemiaminal diastereomers of 4.15.

Additionally, the Porco laboratory noticed that the NH chemical shift was further downfield in the stable hemiaminals 4.15a/b than in the unstable 4.16, which is expected for hydrogen-bonded amide protons.\textsuperscript{306} Hydrogen-deuteration exchange experiments further confirmed this, as the unstable hemiaminal showed rapid exchange with D\textsubscript{2}O while the stable hydrogen-bonded hemiaminal was only half-exchanged after two hours. As hydrogen-bonding networks have been found to influence the bioactive conformation in several natural products such as bryostatin, spongistatin and bafilomycin A, Porco and co-workers hypothesized that it may play a role in the cytotoxicity of zampanolide.
4.3.2 Diaz: Bioactive Conformation of Dactylolide

In 2012, Field and co-workers used two different NMR-based methods, STD (saturation transfer difference)\textsuperscript{307} and TR-NOESY (transferred nuclear Overhauser effect)\textsuperscript{308} spectroscopy, in their attempt to determine the bound structure of zampanolide.\textsuperscript{293} As these methods are reliant on ligands that exchange rapidly with their protein target, they used dactylolide for this experimental work owing to the fact that its covalent binding is significantly slower than that of zampanolide.

The STD-NMR experiment works by measuring intensity differences between a selectively saturated ligand-protein spectrum and an unsaturated one; by irradiating protein signals between 0 and -1 ppm full protein saturation is achieved via spin-diffusion effects. The ligand that binds to the receptor will receive magnetization transfer to its hydrogens through the nuclear Overhauser effect, and as a result should be the only signals to appear on the difference spectrum. Using this experiment, Field and co-workers found that all of dactylolide’s macrocyclic protons were affected when tested with both polymerized microtubules and unassembled tubulin, suggesting an extensive interaction of the molecule with the binding site.
In contrast, a TR-NOESY spectrum is far more useful when attempting to divine a ligand’s bound conformation. This method relies on the difference in correlation time between a free ligand and a bound one; free ligands in solution exhibit small correlation times and small positive NOEs, but when they bind their protein target they become part of a larger macromolecule, which displays longer correlation times and strong negative NOEs. As such, a bound ligand will display characteristic negative NOEs that should correlate to its bound conformation. Field and co-workers observed several key negative NOEs between remote macrolide protons in both forms of tubulin. To match this to a structure, they explored dactylolide’s conformational space by performing a 20,000-step Monte Carlo simulation with the OPLS2005 force field and the GB/SA solvent model in MacroModel. From the resulting library, they selected a single structure which they felt best matched the observed TR-NOESY data. For additional support, they asserted that
ROESY experiments of free dactylolide in buffer indicated that the conformation in water was essentially identical.

![TR-NOESY NOEs between bound dactylolide protons](image)

**Figure 4.5 – TR-NOESY derived conformation of dactylolide bound to tubulin. Figure taken from Field *et al.*, *Chem Biol.*, 2012, 19, 686.**

As shown in Figure 4.5, their proposed conformation is relatively flat. Strong NOE correlations between H15-C17Me and C17Me-H19 indicate that the western fragment of the molecule adopts a conformation that minimizes allylic strain between the C17Me and the C15 stereogenic center. Correlation between H2 and H4 suggest that the northern diene adopts an *s-trans* configuration, and a number of medium and weak TR-NOESY interactions were used to round out the general shape.

4.4 Solution Conformational Analysis of Dactylolide

While the work done by Field and co-workers is compelling, they neglected to use coupling constants as part of their conformational analysis and did not consider any potential flexibility. With the relative dearth of stereocenters and the considerable amount
of unsaturation, the macrolide ring should adopt several additional conformations aside from their proposed flat one. As we believe that having a thorough knowledge of a flexible polyketide’s conformational preferences is a key step in the design of bioactive analogues, our lab embarked on a program to determine the solution conformation distributions of zampanolide and dactylolide through a combination of molecular modeling, high-field NMR experiments and DISCON analysis. 309

4.4.1 Molecular Modeling of Dactylolide

Although zampanolide is the more promising of the polyketide pair, the hemiaminal side chain posed several issues for computational evaluation. As with neopeltolide, initial attempts to model the full molecule were complicated by the program treating conformational changes to the side chain as entirely new structures, which led to an alarming number of structures (>10,000) being found in a 50,000-step Monte Carlo search while limiting the overall number of macrolide conformations. As zampanolide is also more difficult to access synthetically and potentially harder to evaluate via spectral analyses, we instead focused on the (−)-dactylolide (4.14) yielded by our synthetic route for our conformational studies.

Identifying dactylolide’s conformational families began by using our conformational search methodology. MacroModel was used to perform 50,000-step Monte Carlo searches with the MM3*, MMFF*, and OPLS2005 force fields, respectively, with the GB/SA solvation model for water. The structures from the three searches were combined into a single library, and after redundant conformations were eliminated the structures that fell within a 20 kJ/mol window of the global minimum were saved to create a single library.
Interestingly, the vast majority of this library originated from the MM3* search, with a small handful of structures coming from the OPLS2005 search and virtually the entirety of the MMFF* structures being marked as redundant. Similar results were seen in our lab’s conformational analysis of peloruside A, with the OPLS-2001 search producing significantly different results from those derived from MM2 and MM3. This is not entirely surprising, as Class 1 OPLS fields were designed primarily for proteins and lack the additional cross-terms found in the MMX fields.

The flexible regions of the macrolide core were identified using our polar map methodology, with backbone torsional angles plotted onto polar coordinate graphs as mentioned previously. The resulting graphs reveal a mixture of flexible and rigid sections throughout the macrolide core.

Figure 4.6 – Polar maps of dactylolide derived from combined MM3*, MMFF* and OPLS-2005 Monte Carlo conformational searches.
4.4.2 1D and 2D NMR Experiments

While computational analysis can provide valuable insight into molecular flexibility and theoretical conformer populations, conformational analysis ultimately depends on experimental data to narrow the results of a conformational search into conformational families. Unfortunately, prior NMR data for dactylolide was insufficient for our purposes, as a number of methylene protons had yet to be distinguished and several peaks had not been fully deconvoluted to their component coupling constants. Though we considered using D$_2$O as our NMR solvent, DMSO-d$_6$ was ultimately used to best mimic a polar environment while keeping the C19 aldehyde intact. Gratifyingly, high field (600 MHz) $^1$H NMR experiments with our synthetic dactylolide in DMSO-d$_6$ allowed for the identification of a nearly complete set of coupling constants, in particular the previously unresolved C10 protons. Additionally, these values matched well with the coupling constants extracted from natural zampanolide in DMSO-d$_6$, indicating that the solution conformations of dactylolide should also be shared by zampanolide.

In addition to our 1D work, a 2D correlation spectroscopy (COSY) experiment helped with the assignment of dactylolide’s non-overlapping protons. Finally, a ROESY experiment run at an optimum mixing time of 800 ms yielded 39 cross-peaks. Using the C10 methylene protons as a reference, these peaks were normalized and integrated using the equation found in Section 2.6.4 to yield quantitative peak volumes.

4.4.3 DISCON Identification of Dactylolide Solution Conformations

The flexibility observed in the polar maps suggested that dactylolide’s macrolide core possessed several major conformations in solution. To deconvolute the time-averaged NMR data into individual conformers, we employed the DISCON software...
package as described previously. Of the 39 NOE peaks obtained from the ROESY experiment, we selected 11 for use with DISCON along with 6 of the 14 observed $^3J_{H\text{-}H}$ coupling constants based on the clustering statistics dendogram from preliminary runs (Figure 4.7).

DISCON analysis started by determining the optimum number of clusters. Repeated analysis at different clustering levels found that the RMS clustering fit to the NMR experimental data would steadily drop as cluster levels increased, which is consistent with overfitting. However, the RMS and total error roughly levelled out around cluster level 10, suggesting that higher clustering levels would not significantly improve the overall fit to the data. At this level, the solution RMS fit to the NMR data (0.09) was lower than any individual structure (next highest of 0.13). Repeated DISCON runs with different cluster members at this level eventually led to the identification of three representative conformational families that consistently scored higher than the others, of which the lowest energy structures were selected for final visual analysis.
Interestingly, the two major families showed similar conformational profiles save for the northern fragment, which changed the overall shape from hooked (conformer A) to flat (conformer B). The minor family maintains the same conformation across the pyran ring and the trisubstituted olefin, but dihedral changes in the C5-C7 region cause the rest of the structure to twist itself outward (conformer C). Representatives from the
other clusters never exceeded more than 5% contribution to the NMR data fit in DISCON, and were not considered significant to the ensemble.

4.4.4 Conformational Analysis of the Northern Fragment

Computational analysis of the C1-C7 region of dactylolide reveals a fragment that has internal rigidity yet is flexible with respect to the rest of the molecule. The lactone will flip freely to change its orientation by 180° as shown in the 18-19-O-1 and O-1-2-3 polar maps, aligning itself with the diene to maintain orbital overlap in both instances. The diene itself contributes the most to the fragment’s rigidity, preferring to exist in an s-trans conjugated system to both maximize orbital overlap and minimize steric interactions in the C2-C5 section. This orientation was confirmed by a strong H2-H4 cross peak on the ROESY spectrum and an 11.2 Hz $^3J_{H3-H4}$ coupling constant. The majority of the region’s flexibility comes in the C5-C7 region. The 4-5-6-7 torsion adopts two primary dihedral angles of 120° and −120° as a consequence of A$_{1,2}$-strain from the C4-C5 trisubstituted olefin, which explains the conformational family represented by conformer C. ROESY cross peaks between the C6 methylene protons and H3 confirm these conformational preferences.
4.4.5 Conformational Analysis of the Eastern Fragment

The C7-C11 region of dactylolide computationally appears to be rather flexible based on the 6-7-8-9, 8-9-10-11 and 9-10-11-12 polar maps. In particular, the enone moiety rotates on a plane perpendicular to the pyran ring, adopting both s-trans (conformers A and B) and s-cis (conformer C) orientations. As this region contained fairly small ROESY cross peaks, our newly-extracted coupling constants provided the majority of experimental data for our analysis. Time-averaged $^3J_{H9-H10A}$ and $^3J_{H9-H10B}$ coupling constant values of 6.9 and 7.5 Hz supported the observed rotation about the C9-C10 bond, which would be consistent with the averaging of the calculated coupling constant values for conformers A and B. This torsional preference is also a consequence
of A-strain, and the low-energy eclipsed conformations mirror those displayed by 1-butene.

![Conformer A](image1.png)

![Conformer B](image2.png)

Figure 4.10 – Conformational analysis of the C7-C11 enone region.

4.4.6 Conformational Analysis of the Western Fragment

The C15-C17 region of dactylolide appears to have a certain measure of flexibility based on the results of the conformational searches. However, the structures that emerged from DISCON were decidedly rigid, with a strong cross peak between C17Me and the H15 axial proton as well as a strongly anti $^3J_{H18-H19}$ coupling constant of 11.0 Hz. This conformational preference reflects the minimization of allylic strain between the C17 methyl group and the C15 stereocenter, with H15 assuming an eclipsed orientation with C17Me in order to alleviate $A^{1,3}$-strain. A closer examination of the 14-
15-16-17 polar map reveals that conformations possessing a dihedral angle that would introduce allylic strain are noticeably higher in energy.

Figure 4.11 – Conformational analysis of the C15-C19 region.

For the C17-C19 region the polar maps show a reasonable degree of rigidity, with the 16-17-18-19 map primarily showing adoption of only one angle and the 17-18-19-O one showing two. A strong cross peak between H16 and H18A establishes that the latter proton is strongly anti to the C17Me, which would help minimize A\(^{1,2}\)-strain in the C17-C18 torsion. The C18-C19 torsion adopts both 60° and −60° angles, which correspond to two conformations where the macrolide backbone is either anti or gauche to the C20
aldehyde. A large $^{3}J_{\text{H18A-H19}}$ coupling constant of 11.2 Hz confirms that the former is the preferred conformation, which can be rationalized by the minimization of sterically unfavorable gauche interactions.

![17-18-19-O](image)

**Conformation seen in solution**

Figure 4.12 – Possible orientations and observed dihedral angle of the 17-18-19-O torsion.

4.4.7 Comparison of DISCON-derived Solution Conformations to Tubulin-Bound Zampanolide

One of our major structures, conformer B, is nearly an identical match to the TR-NOESY structure proposed by Field and co-workers,\textsuperscript{293} which is unsurprising given the majority of the strong NOE cross peaks they observed were within relatively rigid regions. The more interesting resemblance is between our other major structure, conformer A, and the tubulin-bound conformation of zampanolide, which was published shortly after we identified dactylolide’s major solution conformations. The majority of both macrolides was found to be nearly identical with the exception of the C7-C10 enone
region, which would be expected given the change in C9 hybridization following the Michael addition with H229 in the taxane pocket.

Figure 4.13 – Overlay of major solution conformer A (green) with the tubulin-bound structure of zampanolide (orange) bound to H229 (gray).

However, the conformation of the enone region prior to the formation of the covalent bond was unexpected. The stereochemistry of C9 in the bound complex is inconsistent with the s-trans enone that we observed in the two major solution conformation families, and would instead require the s-cis orientation seen in the minor family to prevent massive steric interactions in the pre-bound complex. As the earlier Flutax-2 displacement experiments demonstrated that both dactylolide and zampanolide form a non-covalent interaction with the taxane site prior to forming the covalent one, this suggests that the macrolide retains considerable flexibility even when non-covalently bound.
4.5 Design of a Simplified Analogue: 17-Desmethyl Dactylolide

The co-crystal structure of zampanolide bound to tubulin raises several interesting questions about the connection between its macrolide conformation and activity. As opposed to epothilone, which conforms nicely to the taxane pocket, zampanolide’s macrolide has minimal significant contact with its binding site, being oriented nearly 90° relative to epothilone and effectively surrounded by water molecules. As zampanolide appears to be held in the taxane pocket largely on the strength of its side chain and the covalent bond with H229, the macrolide could be little more than a glorified spacer between a covalent anchor and the active hydrogen bonds with tubulin. Given the difficulty in synthesizing zampanolide and the limited amount of material available for biological testing, we envisioned a simplified macrolide that would test the role of conformation in zampanolide’s biological activity while shortening the number of necessary synthetic steps.

4.5.1 Previous Analogues: Altmann 2012

Despite several reports of formal and total syntheses of dactylolide and zampanolide, SAR studies on either molecule have been fairly limited. The only significant efforts to date have been performed by Altmann and co-workers in their attempt to identify the pharmacophoric elements required for microtubule-stabilizing and antiproliferative activity. Their work so far has focused on exploration of the zampanolide side chain and the pyran region of the shared macrolide. As covered earlier, the C20 aldehyde of dactylolide is highly susceptible to attack from protic nucleophiles such as methanol. This susceptibility extends to water as well, as Hoye and co-workers found that the aldehyde undergoes rapid hydration to form a C20
geminal diol. From a bioactivity standpoint, the resulting hydrate likely mimics zampanolide’s C20 alcohol and makes the corresponding hydrogen bond with T274 in the taxane site. Altmann’s first analogue, a late-stage alcohol intermediate 4.17 in their dactylolide synthetic route, demonstrates this nicely by displaying antiproliferative activity on par with that of dactylolide (IC$_{50}$ = 127 ± 2.9 nM, A549), confirming that the aldehyde does not play a role in tubulin binding. Accordingly, methylation of this alcohol resulted in a 10-fold loss of cytotoxicity due to the loss of the hydrogen bond donor (IC$_{50}$ = 1072 ± 103 nM, A549). Altmann and co-workers next attempted to enhance dactylolide’s water solubility by forming the carboxylate 4.19, but the near loss of cytotoxicity (IC$_{50}$ = 9732 ± 260 nM, A549) by this analogue led them to conclude that it possessed poor cell penetration.
The stereoselective incorporation of the $N$-acyl hemiaminal side chain to convert dactylolide to zampanolide represents a significant challenge and one of the major reasons for the lack of synthetic zampanolide. Altmann and co-workers decided to investigate the necessity of this sensitive functionality by replacing it with n-hexylamine. Interestingly, the resulting amide 4.20 displayed moderate cytotoxicity ($IC_{50} = 973 \pm 90$ nM, A549) against several cancer cell lines, which is surprising given the increase in conformational flexibility with the loss of the diene and the apparent lack of both key hydrogen bonding moieties with the taxane pocket.
The Altmann group next shifted its focus to the functional role of the pyran ring and the possible simplification of the macrolide. Their first analogue was designed to probe the role of the C13 exocyclic alkene by targeting desmethylene analogues of dactylolide. The resulting compounds 4.21 and 4.22 both displayed nearly equipotent antiproliferative activity compared to their precursors (IC$_{50}$ = 189 ± 19.3 nM and 149 ± 12.8 nM, respectively, A549), strongly suggesting that the alkene plays little role in affecting biological activity. Observation of the zampanolide co-crystal structure appears to confirm this, with the moiety occupying a shallow hydrophobic pocket well away from the M-loop and the covalent binding interaction.

Altmann and co-workers decided to follow up on these promising results by excising the tetrahydropyran subunit in its entirety, which would yield a vastly simplified macrolide containing only one stereocenter. They prepared three compounds based on this scaffold, alcohol 4.23, aldehyde 4.24, and the full desTHP-zampanolide 4.25, though their methodology to install the side chain provided the latter compound as a 1.6/1 diastereomeric mixture. While 4.23 and 4.24 showed a considerable loss in activity compared to their pyran-containing forebears (IC$_{50}$ = 2378 ± 70 nM and 3921 ± 216 nM, respectively, A549), 4.25 displayed antiproliferative activity comparable to that of natural dactylolide despite its lack of purity and represented a nearly 20-fold increase over the other desTHP analogues (IC$_{50}$ = 165 ± 13 nM, MCF7). These results further support the crucial role the hemiaminal side chain plays in zampanolide’s biological activity.
4.5.2 Simplification of the C16-C17 Trisubstituted Olefin

The potency displayed by 4.25 raises interesting questions on the connection between zampanolide’s macrocyclic conformation and its biological activity. The removal of the THP ring of zampanolide also brings the loss of rigidifying A$^{1,3}$-strain, which was present in all three of our observed solution conformational families. While all three desTHP analogues are considerably less active than their parent compounds, the nanomolar activity of this analogue suggests that the hemiaminal side chain plays a considerably larger role than the macrolide.
The $E$-trisubstituted alkene of zampanolide has historically represented a small roadblock for total synthesis efforts, particularly those that would be amenable to high overall yields. Methods used to install this group as part of a convergent pathway possess notable drawbacks, whether in the form of requiring unstable vinyl iodide intermediates\textsuperscript{311} or olefination reactions with poor $E/Z$-selectivity and control.\textsuperscript{312} Our own synthetic route was able to incorporate this functionality with high selectivity, but incorporating the C19 stereocenter required several additional manipulations that curbed the overall yield of the desired fragment.\textsuperscript{301} As the C17 methyl group appears to play no other role than to induce A\textsuperscript{1,3}-strain in both zampanolide and dactylolide, we envisioned that its removal would yield a simplified analogue that could be produced with a shorter step count in higher yields.

4.5.3 Conformational Analysis of Simplified Dactylolide Analogue

As mentioned previously, one of the defining characteristics of dactylolide’s conformational families is the highly rigid C15-C19 region of the macrocycle. Our conformational analysis of this region found that the two involved torsions, 14-15-16-17 and 16-17-18-19, respectively adopt major dihedral angles of $-120^\circ$ and $120^\circ$ as a consequence of allylic strain. As the removal of the C17 methyl group responsible for the strain on both torsions should significantly affect the conformational flexibility of this region, we decided to investigate any new theoretical conformational families and if they would significantly change the overall conformer population.
Building off the work done by Altmann and co-workers, we decided to use their 13-desmethylene analogue 4.21 as a starting point for our own analogue work. Although we decided to target the macrocyclic alcohol 4.26 synthetically, for modeling purposes we used the corresponding aldehyde 4.27 in order to minimize changes from our dactylolide studies. Following our previous modeling protocol, we ran 50,000-step Monte Carlo searches using the MM3* force field and the GB/SA water solvation model on structure 4.27. Structures that fell within a 20 kJ/mol window of the global minimum were saved and screened for redundant conformers to yield the final library, which was used to generate polar maps of the macrolide torsions.

Unsurprisingly, many of these polar maps were almost identical to the ones found in our dactylolide searches. As molecular mechanics force fields are relatively limited in scope, local acyclic interactions such as those detailed in Chapter 5 will largely dictate the resulting conformational preferences and families. However, the C15-C19 region saw a drastic increase in flexibility with the removal of the C17 methyl group. From the library, we identified two low energy conformers, D and E, which were representative of
the observed conformational preferences. Conformer D is virtually identical to the
dactylolide solution conformer A and the zampanolide crystal structure, the former of
which also being the global minimum of the dactylolide conformational searches.
Conformer E, on the other hand, displayed an entirely new C15-C19 region.

Figure 4.17 – Low-energy conformations observed from a Monte
Carlo conformational search of 4.27.

As shown in Figure 4.18, new dihedral angles for the 14-15-16-17 and 16-17-18-19
torsions appear at 0° and −120°, respectively. In our dactylolide solution conformation
families, the 14-15-16-17 torsion exclusively adopts a −120° angle that eclipses the C17
methyl group with the C15 hydrogen rather than the C14 alkyl protons or the
tetrahydropyran oxygen, thus minimizing A\textsuperscript{1,3}-strain. Replacing the C17 methyl with a
proton still favors this torsion, but also allows for the C14 eclipsed conformation to
become viable. The reason for this can be seen in 4-methyl-2-pentene model systems
(Figure 5.4, Chapter 5); the allylic strain interaction only incurs a 0.73 kcal/mol increase
in energy in the E system, while the Z allylic strain is in excess of 4 kcal/mol.
For the 16-17-18-19 torsional angle, our dactylolide conformations showed preference for a 120° torsional angle, which has one of the C18 hydrogens eclipsing the C16 vinylic proton in order to minimize $A^{1,2}$-strain. While the alternate $-120^\circ$ torsion would also minimize $A^{1,2}$-strain, adopting this angle in dactylolide would incur the previously mentioned $A^{1,3}$-strain in the 14-15-16-17 torsion. With the C17 methyl removed, the 16-17-18-19 torsion is able to adopt the alternate conformation without incurring significant energy penalty.

Figure 4.18 – Observed conformational preferences of conformers D and E in the changed polar maps.
Interestingly, conformer E displays an alternate conformational preference for the 17-18-19-O dihedral angle, favoring the −60° angle rather than the 60° observed in dactylolide. The 17-18-19-O polar map is identical for both compounds, and found that both conformational families should be likely. However, our solution NMR work on dactylolide found this to be a rather rigid torsion based on the strong H18-H19 coupling constant of 11.0 Hz, which suggests that the 17-desmethyl analogue would resemble conformer D in this area. The remainder of conformer E bears a strong resemblance to the twist conformation observed in dactylolide solution conformer C.

4.5.4 Synthesis of 17-desmethyl-13-desmethylened Dactylolide

![Scheme 4.6 – Retrosynthesis of simplified analogue 4.26.](image)

To test our theory that the relatively similar conformational profile of our dactylolide analogue predicts biological activity, Matt Wilson began designing the synthesis of 4.26.\(^{301}\) As the analogue is a simplification of Altmann’s the retrosynthesis targets intermediate 4.28, which is similar to the one from their route (Scheme 4.6). One of the key stumbling blocks in our synthetic route to zampanolide, and one of the driving
forces behind the design of this analogue, came from the installation of the trisubstituted olefin of zampanolide at C16. The original scheme, which envisioned installing both the olefin geometry and the C19 stereocenter simultaneously, instead led to a 1.1:1 $E/Z$ isomeric mixture (Scheme 4.7). As a result, the olefin geometry instead had to be installed using the much simpler methacrolein as part of the tandem cross-metathesis/HWE reaction outlined earlier, with several additional steps needed to reach the desired intermediate.

![Scheme 4.7](image)

Scheme 4.7 – Initial (failed) plan to install trisubstituted olefin in dactylolide/zampanolide core.

Conversely, synthesis of 4.28 proceeded with far less difficulty. The loss of the exocyclic C13 olefin meant that the tetrahydropyran fragment 4.30 could be synthesized in only three steps, capped off using a Tsuji-Trost cyclization. This fragment was then coupled directly to an enantiopure homoallylic alcohol 4.31 using HG-II to yield the key intermediate. The C17Me group was preventing the use of this same cross-metathesis for direct installment of the C16-C20 fragment with any olefin selectivity, so gratifyingly the switch to a disubstituted alkene allowed for this reaction to proceed with high yield and $E$-selectivity. The remainder of the synthesis used Altmann’s route, coupling the C1-C8 fragment 4.29 using a Yamaguchi esterification and a Horner-Wadsworth-Emmons
olefination to complete the macrocycle. Overall, the synthesis of this analogue required only 15 steps in the longest linear sequence, as opposed to the 26 required for our lab’s synthesis of dactylolide. Additionally, many of these intermediates were achieved in significant quantities using known scalable routes, which bodes well for future studies should the analogue prove biologically active.

Scheme 4.8 – Synthetic route to target intermediate 4.28.
5.1 Purpose

The purpose of this chapter is to present our efforts to determine whether the conformational preferences of common polyketide structural units observed in solution and computational molecular modeling could be affected by crystal packing forces in a solid-state environment. An examination of the electron-crystallography co-crystal structure of epothilone A bound to tubulin, which featured high-energy syn-pentane interactions in the former, will be reviewed, followed by an exploration of some possible reasons for ligands to adopt conformations that differ from their observed solution energy minima. Finally, our search of crystallographic databases for some of the polyketide structural features outlined in previous chapters will be presented, along with our rationale for deviations from otherwise expected conformations.

5.2 The Journey to the Bound Conformation of Epothilone

As mentioned in Section 1.6.1, the ongoing efforts of research laboratories to determine the bound conformation of the epothilones and its relation to paclitaxel in the early 21st century were shaken by the publication of an EC-derived structure in 2004.161 This structural orientation was markedly different from previously reported solution, TR-NOESY- and analogue-derived conformations, with the C11-C15 region adopting the
minor solution conformational orientation observed by the Taylor laboratory.\textsuperscript{156} Most strikingly, however, was the C1-C9 backbone region, which now possessed two atypical high-energy syn-pentane interactions as part of a previously unreported conformation (Figure 5.1). Due to the limited resolution of the binding pocket (3.8Å), Nettles and co-workers derived their epothilone A structure with the help of a NAMFIS ensemble. They used this conformation and its purported binding pose to explain known aspects of epothilone’s SAR, propose an apparent shared pharmacophore with paclitaxel and rationalize some of the known resistance mutation in β-tubulin.

Figure 5.1 – EC-derived tubulin-bound conformation with observed syn-pentane interactions (PDB 1TVK, top left) compared to the solution/X-ray structure of epothilone A (top right), the CytP450K bound conformation (bottom left), and the TR-NOESY structure in unassembled tubulin (bottom right).
However, attempts to design analogues that would verify their conformation as bioactive were generally unsuccessful. The Snyder laboratory first restrained the epothilone A macrolide by creating a bridge between C6 and C8, which would lock in the \textit{syn}-pentane moiety observed in their EC structure and replicate the overall backbone conformation according to their docking work.\textsuperscript{314} Preliminary evaluation of this bridged compound \textit{5.1} found that it was only weakly active (IC\textsubscript{50} = 8.5 μM, A2780 cells), representing a potency loss of nearly 3900-fold. Soon thereafter, work by the Kingston laboratory focused on epothilone D, creating a bridge between C4 and C12 to restrict the macrolide.\textsuperscript{315} Among their reported analogues two in particular stand out with respect to the effects of the conformational restriction, \textit{5.2} and its ring-closing metathesis precursor \textit{5.3}. \textit{5.2} was predicted to match the EC-derived conformation, but was found to be nearly 100-fold less potent against A2780 and PC3 cells when compared to epothilone D. \textit{5.3}, on the other hand, was only 2 to 5-fold less active than the parent compound, further suggesting that the EC-derived conformation was detrimental to biological activity. In contrast, analogues prepared by the Taylor laboratory demonstrated that conformations which reflected epothilone’s major solution conformational preferences retained activity, as covered previously.\textsuperscript{156} Curiously, Nettles et al attempted to explain the biological activity of the relevant C14-substituted epimers in their work, but reversed the reported SAR in favor of their reported structure in their model figure despite correctly identifying them in the text (as noted by Heinz et al in their review of said work\textsuperscript{316}). In 2013, Prota and co-workers resolved the debate by publishing a high-resolution X-ray crystal structure of epothilone A bound to tubulin that conclusively identified its bioactive conformation’s resemblance to the solution structure, as seen in Figure 1.32.\textsuperscript{165}
Figure 5.2 – Epothilone analogues designed to be conformationally-restricted to the Nettles et al EC structure.

5.3 The Effects of Environment on Ligand Binding

Given the considerable evidence against the EC-derived structure of epothilone A, it begs the question as to why this was considered the definitive bioactive conformation for nearly a decade. Crystal structures of epothilone B and D in complex with cytochrome P450epoK,\textsuperscript{317} which catalyzes the last step in the biosynthesis of epothilone A and B, had been previously found to adopt conformations resembling that of Taylor’s minor solution structure,\textsuperscript{145} but still were quite distinct from the EC structure. The original work by Nettles and co-workers offers little justification for their structure outside of the pharmacophore argument, focusing instead on the interactions it would make with the β-tubulin binding site. In their highlight of the Science paper, Heinz et al attempt to rationalize the differences in reported conformations as a consequence of the artificial environments in which they were collected.\textsuperscript{316} The EC structure of Nettles et al...
was bound to α,β-tubulin in zinc-stabilized sheets, the TR-NOESY structure of Carlomagno et al was obtained using unassembled tubulin, and the solo solution and X-ray crystallographic structure could be a result of solvation or crystal packing forces, respectively.

The assertion that a change of environment could result in different bioactive conformations of epothilone is particularly puzzling. The highly-substituted structure of epothilone makes it unlikely that it would preferentially adopt high energy syn-pentane interactions in a bioactive conformation, particularly when it has been demonstrated to possess more than one conformation that avoids them. Moreso, work by the Danishefsky laboratory found that removal of the C8 methyl involved in one of these interactions led to a 200-fold reduction in biological activity, which is counter-intuitive as this should have lowered the energy of the same macrolide conformation.318

5.3.1 Ligand Binding Strain

On the surface, the idea that a molecule like epothilone could adopt a bound conformation higher in energy than its base solution or solid state structure is not particularly novel. The induced fit model proposes that both a ligand substrate and its target binding site will induce conformational changes in one another as they bind to create the optimum configuration.319 As long as any increase in energy created as part of this arrangement due to strain is offset by the overall complex lowering in energy, the binding remains thermodynamically favorable. In 2004, Perola and Charifson reported an extensive study of ligand reorganization upon binding for 150 crystal structures of pharmaceutically-relevant protein-ligand complexes.320 They found that not only did 60% of the ligands not bind in a local minimum conformation, but that at least 10% of their
ligands would bind with strain energies in excess of 9 kcal/mol, regardless of the method they employed. In this context, the additional 1-2 kcal/mol imparted by something like a syn-pentane interaction does not seem particularly onerous.

Figure 5.3 – Sample of Perola’s analysis applied to PDB 1EZQ. Global minimum conformation (left) is ~100 kJ/mol lower in MMFF potential energy than bound conformation (right).

However, Perola and Charifson’s study does not necessarily reflect the conformational changes that compounds like epothilone would experience upon binding. Almost all of their ligands are hydrophobic, linear and heavily feature aromatic moieties, as opposed to the highly-substituted macrolide ring possessed by epothilones and similar polyketides (Figure 5.3). In fact, they make a special point to mention that a significant contributor to the strain energies experienced by these ligands during binding was the unfolding of their hydrophobic functionalities to expose them to contact with protein
residues. As a result, strain from acyclic forces such as syn-pentane interactions or gauche effects were not well-represented in their analysis.

5.4 Preferred Solid-State Conformations for Common Polyketide Features

Another intriguing facet of epothilone is the similarity of its solid state conformation to its solution conformation, which in turn is highly similar to the bioactive conformation. The general assumption of crystallography is that molecular structures will concentrate in low-energy regions of conformational space in the solid state as they would in the gas phase, effectively suggesting that crystallized molecules should correspond to an intramolecular energy minimum, save for a very small number of compound classes such as substituted and unsubstituted biphenyls. However, research has demonstrated that flexible molecules can have a solid-state conformation that deviates considerably from its solution one due to the packing forces present within the crystalline environment, such as the case of N-methylacetamide and adrenaline.

The 15 years of debate over the bioactive conformation of epothilone piqued our curiosity as to whether the conformational control features of macrocyclic polyketides could be affected by crystal packing forces. In particular, we wanted to gauge the types of features that could be affected by a crystalline environment, and whether they could be induced to adopt high-energy conformations. Though functional groups that could participate in intramolecular hydrogen bonding are the most likely to be altered, van der Waals interactions could theoretically affect hydrophobic areas as well. As such, we decided to search available crystallographic databases for the conformational features outlined previously in Figure 1.4, to investigate if and how their preferred angles deviate from those observed via experiment or molecular mechanics simulations.
5.4.1 Methodology

For our searches, we decided to use the Cambridge Structural Database (CSD), which is overseen by the Cambridge Crystallographic Data Centre (CCDC), as it represents the most significant available repository of small-molecule organic and metal-organic crystal structures with over 800,000 entries. The database was searched using the ConQuest software available as part of the CSD System, which can scan the entire library locally using Boolean search terms to eliminate unwanted structures. In general, we try to limit our searches to the features as presented in Figure 1.4, since they represent the moieties at their simplest and experience minimal deviation as a result of other functional group interactions. We also made a point to eliminate functionalities that were part of rings smaller than 9 members, as this would eliminate ring strain effects on conformation. ConQuest searches were exported to Mercury (also part of the CSDS), which was used to visualize crystal structures, generate polar histograms, or export torsional data to Excel for any additional manipulations needed.

5.4.2 Structural Features Affecting One Bond

5.4.2.1 Allylic 1,3-Strain

Allylic 1,3-strain, also referred to as $A^{1,3}$-strain, results from the interaction between a substituent on one end of a Z-olefin with an allylic substituent on the other end. If the substituents are large enough in size, they can sterically interfere with one another to the degree that one conformation will be favored over the other. Allylic strain was first recognized in the literature by Johnson and Malhotra in 1965, which they used to explain conformational changes in cyclohexane rings containing exocyclic olefins.\textsuperscript{325}
However, this effect may be seen as well in simple olefins. *Ab initio* calculations performed on 3-methyl-1-butene (5.4) show that the energy minima are represented by conformations 5.4a and 5.4b, with 5.4c at least 2 kcal/mol above the global minima. Though 5.4a would be favored at equilibrium, the relatively low energy difference means that 5.4b is not significantly disfavored. The situation changes when a larger group is substituted onto the olefin, as in the case of 4-methyl-2-pentene (5.5). Now, the conformational equilibrium strongly favors 5.5a, because conformation 5.5b is destabilized by steric interactions between the two substituents. As this work has demonstrated, the energy penalty incurred via A1,3-strain is enough to significantly influence conformational distributions, as seen in the C14-substituted epothilones and zampanolide.

Figure 5.4 – Energy changes arising as a result of torsional changes in 5.4 and 5.5.
Searching the CSD for this functionality found 410 instances, which were analyzed using Mercury. The average torsional angle observed for $\theta$ was $-1.8^\circ$, which would place the hydrogen fully eclipsed with the olefinic methyl group as expected. Additionally, we see a standard deviation of $22.3^\circ$ with a corresponding drop-off in population as the angle deviates from the ideal value. This also fits what is known about $A^{1,3}$-strain, as changes in the dihedral angle $\theta$ are possible at a cost of $<1\text{kcal/mol}$. From this data, we can conclude that $A^{1,3}$-strain appears to be unaffected by potential forces within a solid-state environment.

Figure 5.5 – Polar histogram for the dihedral angle $\theta$ representative of $A^{1,3}$-strain.
5.4.2.2 1,2-Diols – The Gauche Effect

In the classic case of staggered $n$-butane, the two terminal methyl groups will adopt either an *anti* relationship or a *gauche* relationship. Normally, the *gauche* rotamer will be less stable than the *anti* one by approximately 0.9 kcal/mol, for both steric and electronic reasons. The Gauche effect refers to any situation where this preference becomes reversed as a consequence of electronegative functional groups replacing the methyl groups, as in the case of 1,2-difluoroethane. Here, the donation of electron density from a hydrogen atom to the C-F $\sigma^*$ anti-bonding orbital acts as a source of stabilization for the gauche isomer (Figure 5.6). Conversely, the *anti* isomer becomes destabilized by approximately 0.7 kcal/mol over the *gauche* form, as the increased electronegativity of fluorine makes the C-F $\sigma$ orbital a poor electron donor.\(^{327}\) This effect is also observed in 1,2-dimethoxyethane, though to a lesser degree ($\Delta E = 0.14$ kcal/mol).\(^ {328}\) In macrocyclic polyketides, we observe this effect through vicinal electronegative functional groups, usually in the form of diols (e.g. erythromycin, hypothemycin, alchivemycin A) or a combination of hydroxyl and methoxy groups (e.g. tedanolide, geldanamycin, rapamycin, herboxidiene).
Initially we searched the CSD for vicinal diols, which found 2500 instances. A polar histogram of the O-C-C-O dihedral angle found that a significant population of these instances fell at 60° or -60°, which would reflect the Gauche effect. However, approximately a third instead fell around 180°, which would reflect an anti orientation that should be stereoelectronically disfavored. Cursory examination of some of these anti isomers found that many of them seemed to be smaller and linear, which appeared to be allowing intermolecular hydrogen bonding between molecules in the crystal structure to influence conformation (Figure 5.7). To check this, we limited the search to vicinal diols that were part of a macrocycle (>9 atoms), which cut the library to 202 instances. In this subset the anti rotamer was considerably less populated, representing around 10%. Virtually all of the hydroxyl groups were involved in intermolecular hydrogen bonding of
some kind, and a sizeable number of these *anti* torsions occurred in the middle of a series of hydroxyl groups.

![Figure 5.8 – Observed torsional angles for vicinal diols in non-macrocycles (left) and macrocycles (right).](image)

Finally, we searched the CSD for the hydroxyl-methoxy vicinal combination mentioned previously. This search found 117 instances for both rings and linear structures, with nearly a quarter (28) exhibiting an *anti* O-C-C-O dihedral angle. Limiting this search to macrocycles reduced found only 12 structures, of which 4 possessed the *anti* configuration. Interestingly, the hydroxyl groups of these latter structures were all involved in hydrogen-bonding to adjacent molecules in unit cell crystal packing, while only three of the eight possessing *gauche* configurations had similar participation.
5.4.2.3 Esters/Lactones

Ester functionalities are common features of polyketide natural products. Their cyclic variety, the lactone, is nearly omnipresent among macrocyclic polyketides as a consequence of the cyclization that occurs to separate the growing molecule from its primary PKS pathway. As mentioned in Section 2.6.1, ester functionalities will preferentially adopt an s-cis orientation that aligns a lone pair on the alkyl oxygen anti-periplanar to the C=O bond. This preference can be rationalized through some combination of hyperconjugation, dipole minimization and possible steric interactions, with the s-cis form roughly 3 kcal/mol more stable than the s-trans rotamer depending on solvent and the size of R and R’. Searching the CSD for structures where the s-trans form would not be preferred due to ring strain yielded predictable results, with the vast majority of the 27,000 instances found adopting the s-cis form. The ~100 or so that saw

Figure 5.9 – Observed torsional angles for OH-OMe gauche effect.
deviation from the ideal 180° torsion were mostly due to extraordinary inter- and intramolecular interactions caused by a high concentration of aromatic functionalities or long, saturated hydrocarbon chains.

![Diagram showing torsional preferences for esters and lactones](image)

**Figure 5.10 – Torsional preferences for esters and lactones.**

5.4.2.4 α-Substituted Ketones

The high oxygenation character and substituent count of polyketides also makes α-substituted ketones a common sight among its macrocyclic members, including several already covered in this work (e.g. tedanolide, spongistatin, epothilone). As mentioned previously, ketones will prefer to adopt conformations where the carbonyl bond eclipses with the alkyl group rather than a hydrogen. This allows the alkyl groups on both sides of the carbonyl to be *anti* rather than *gauche*, as seen in the preferred conformation of 2-butanone.\(^{259}\) Increased substitution at the alpha position still prefers an alkyl-eclipsed conformation over a hydrogen-eclipsed one by approximately 1.8 kcal/mol.\(^{329}\)
Figure 5.11 – Preferred conformations of 2-butanone (left) and 3-methyl-2-butanone (right).

Searching the CSD for α-methyl ketones found 385 instances, with the O-C-C-H torsional angle targeted for analysis. *Ab initio* calculations on 3-methyl-2-butanone found that the energy minima of the eclipsed conformation occurred when this dihedral angle was 140°, which was termed “skew”. Interestingly, the observed dihedral angle appeared to prefer 180°, which would instead correspond to an *anti* orientation more in line with the steric effects we previously observed in our searches for A$_{1,3}$-strain. Additionally, approximately 15% of the instances instead adopted the apparently disfavored hydrogen-eclipsed conformation, though examination of the structures revealed that a non-trivial portion appeared to be under the influence of competing structural features (e.g. syn-pentane interactions, intramolecular hydrogen bonding). Given the relatively low energy barrier of rotation, it appears that α-substitution of a ketone with alkyl substituents can play a role in influencing conformation but not a definitive one.
Figure 5.12 – Observed torsional angles for α-methyl ketones. Expected “skew” conformation (top right) vs. major observed anti conformation (bottom right)

Replacing the methyl substituent with a hydroxyl one also yielded surprising results. In our computational study of 2-hydroxy neopeltolide in Section 3.3.2, we postulated that a vinylogous anomeric effect could be responsible for the observed polar map change, which would place the hydroxyl group perpendicular to the carbonyl bond. As such, when we searched the CSD for this functionality we defined the selected torsion as the O-C-C-O dihedral angle, anticipating that we would see ~90° as the preferred angle. However, the resulting 160 instances found that virtually every torsional angle was represented in some fashion, with the primary observed angle near 0°. While this would normally suggest intramolecular hydrogen bonding between the hydroxyl and ketone groups, searching through the structures found this was not always the case. For the alternative conformations, we also were unable to determine any possible correlation between external hydrogen bonding and dihedral angle.
5.4.3 Structural Features Affecting Two Bonds

5.4.3.1 $\textit{Syn}$-Pentane Interactions

$\textit{Syn}$-pentane interaction, also known as pentane interference, is the steric hindrance that occurs in one of the conformations of $n$-pentane. Here, the gauche$^+$ and gauche$^-$ conformation places both methyl groups into close proximity with one another, similar to what would be found in a 1,3-diaxial arrangement on a cyclohexane ring. The resulting steric interaction is comparable to what is observed in $\text{A}^{1,3}$-strain, and as such conformers with a $\textit{syn}$-pentane interaction are higher in energy than an unstrained one by a similar amount (approximately 3–4 kcal/mol).$^{330}$ For this reason, linear hydrocarbon chains will tend to adopt conformations that are free of any such interactions. Similarly, polyketide natural products that possess alternating methyl groups on their backbone,
such as epothilone and discodermolide, will preferentially adopt conformations that avoid syn-pentane interactions, as we have covered previously.

Figure 5.14 – Ramachandran plot for observed torsional angles of cis-1,3-dimethyl moieties.

Investigating this functionality required separating cis-1,3-dimethyl moieties from trans-1,3 ones, as these would present different preferred angles. For the cis structures, we would expect backbone dihedral angle combinations of 60° (ϕ) and 180° (ψ) or 180° (ϕ) and -60° (ψ), as these would represent conformations that eliminate syn-pentane interactions. Searching the CSD found 86 instances of this moiety, of which the vast majority possessed the expected torsional angles when graphed onto a Ramachandran plot (Figure 5.14). Of the 3 instances that adopted torsional angles that would be closer to
syn-pentane generation, one was in a macrocycle subject to ring strain from unsaturation and two were part of longer acyclic hydrocarbon chains. The latter two appear to be the only ones affected by crystal packing forces.

![Torsional angles](image)

**Figure 5.15** – Ramachandran plot for observed torsional angles of *trans*-1,3-dimethyl moieties.

For *trans*-1,3-dimethyl structures, we would expect mirroring dihedral angles for both $\phi$ and $\psi$ of either 60° or 180° to eliminate *syn*-pentane interactions. Searching the CSD for this moiety found a smaller number of instances, though with a tighter distribution around the preferred dihedral angles. Of the marked exceptions, two came from the same compound that possessed a ring heavily constrained by amide bonds. The third (marked in green) came from a synthetic precursor to kermomycin,\(^{331}\) with a
deviated φ angle of 133° but an ideal ψ value of 60°. Interestingly, kendomycin itself was also represented in this CSD search, but with both φ and ψ values very close to ideal values of 180°. In total, though we did not find a significant number of instances compared to some of our other searches, the overwhelming percentage of structures that reflected syn-pentane avoidance with the least-sterically demanding functional groups suggests that crystal packing is highly unlikely to affect these moieties.

5.4.3.2 1,3-Diols

As opposed to the functionalities discussed thus far, 1,3-diols can potentially act as a source of conformational stability through intramolecular hydrogen bonding. 1,3-diols can be found in dictyostatin and polyene antibiotics like amphotericin B, and structures covered elsewhere in this work have demonstrated the potency of internal hydrogen bonds and hydrogen bonding networks as conformational influences. Experiments on cis-cyclohexane-1,3-diol have found that the strength of an intramolecular hydrogen bond between two hydroxyl groups in a syn diaxial arrangement is highly dependent on its environment, with bonds of ~1.6 kcal/mol observed in non-polar solvents and strength generally decreasing as environment polarity increases.\(^{332}\)

Investigating this functionality required a similar approach to the one employed for syn-pentane interactions, separately defining cis- and trans-1,3-diols into their own searches. For the cis variant, we would expect both φ and ψ dihedral angles to be at 180°, which should correspond to an internal hydrogen bond to offset what would otherwise be stericly unfavorable. Searching the CSD for this moiety found 182 instances, but upon making the Ramachandran plot we found that a clear pattern did not emerge. While there was a substantial population at the 180°-180° intersection, a notable number of structures
adopted torsional angles that would instead reflect avoidance of steric interactions akin to what we observed with cis-1,3-dimethyl structures. The only structures that appeared to consistently possess intramolecular hydrogen bonds were larger macrocycles with extensive hydrogen bonding networks, such as chainin or the amphotericin B, where bulky crystal packing would not allow for better intermolecular hydrogen bonds. Additionally, many structures possessed substituents on the carbon in between the two hydroxyl groups (usually a methyl or hydroxyl group), which could also be affecting conformation through gauche interactions.

Figure 5.16 – Ramachandran plot for the observed torsional angles of cis-1,3-diol moieties.
Searching the CSD for *trans*-1,3-diols yielded more instances (346) with a slightly cleaner Ramachandran plot. Here, we effectively observed four possible configurations, two of which represented intramolecular hydrogen bonding ($\phi$:180°/$\psi$: -60° and vice-versa) and two that represented steric avoidance ($\phi$/\psi: 60° or 180°). The vast majority of these instances (278, 80.3%) were located near the 180°-180° intersection, which strongly indicates this functionality tends to avoid steric interactions rather than adopt an intramolecular hydrogen bond. As with the *cis* diols, a significant number of the representative structures possessed substituents on the carbon between the hydroxyl groups that could play a role in conformational preferences. It also bears mentioning that a significant number of the structures found by both the *cis* and *trans* searches were small, linear carbohydrates, which are heavily oxygenated and offer a wide range of possibilities for intermolecular hydrogen bonding.

![Ramachandran plot for *trans*-1,3-diol moieties](image)

Figure 5.17 – Ramachandran plot for the observed torsional angles of *trans*-1,3-diol moieties.
5.4.3.3 β-Hydroxy Ketones

Similar to the 1,3-diols, β-hydroxy ketones can also potentially offer conformational stabilization through internal hydrogen bonding. These functionalities can be found in several polyketides already discussed in this work, such as epothilone, amphidinolide and spongistatin. In particular, our study of the epothilones found that the minor solution conformation adopted a C1-C9 region that placed the C7 hydroxy group within hydrogen bonding range of the C5 ketone, which could account for its presence (see Figure 1.28). Energetic estimates of this moiety in small aliphatic compounds found that the strength of the hydrogen bond can vary wildly depending on carbon substitution, ranging from 1.4 to 13.7 kcal/mol in the gas phase, though most structures without C2 substitution fell around 2-3 kcal/mol.333

With the introduction of an sp² carbon, we decided to measure the two dihedral angles slightly differently to eliminate the need to separate structures based on the chirality of the hydroxyl group. Here, we would expect the ψ angle to be close to 180° to minimize steric interaction between both sides of the carbonyl and the φ O-C-C-C(=O) dihedral angle to be somewhere between 60° and -60° to keep the hydroxyl group close enough for an internal hydrogen bond to be feasible. Searching the CSD for this functionality found 100 instances, specifically excluding substituents on the carbon separating the hydroxyl and carbonyl. Of these instances, approximately half fell into the region that would precipitate hydrogen bonding (marked in red). However, closer examination of many of the individual structures found that the hydroxyl hydrogen frequently appeared to be involved in external hydrogen bonding with other molecules despite being in otherwise optimal position. This suggests that the stabilizing effects of
forming a crystal outweighed the energetic benefits of intramolecular hydrogen bonding, which follows with our observations of the 1,3-diols.

Figure 5.18 – Ramachandran plot for observed torsional angles of β-hydroxy ketone moieties. Ideal hydrogen-bonding regions are outlined in red.

5.5 Conclusions

The trends observed in our searches of the CSD strongly indicate that the influence of a solid-state environment on the conformation of common polyketide motifs is dependent on hydrogen bonding potential and the level of energetic penalty. Each of the moieties that included groups that could act as both hydrogen bond donors and
acceptors (1,2-diols, 1,3-diols and α-hydroxy and β-hydroxy ketones) demonstrated that conformations deviating from expected values were prevalent among crystal structures, with the α-hydroxy ketones providing the full spectrum of torsional angles. Additionally, moieties with lower energy penalties, such as the α-methyl ketones, showed a wide range of highly populated torsions. However, functionalities that are hydrophobic or whose preferences are created by significant strain energy were much better defined, as we saw in the case of syn-pentane generating motifs (A\textsubscript{1,3}-strain, 1,3-dimethyls) and the s-cis esters/lactones. Overall, it appears that the possible stabilization that can be generated through intermolecular hydrogen bonding is sufficient to alter the local conformation to a degree, while high-energy hydrophobic areas less likely to interact with the surrounding environment are far more likely to retain their preferred low-energy conformations.
6.1 Neopeltolide

Our conformational analysis of neopeltolide indicates that the macrolide core likely adopts only a single conformational family in solution, with the s-cis lactone moiety able to rotate about the C1-C2 and O-C13 bonds. Conformational analysis of the reported diastereomeric macrolide analogues found that each displayed noticeable differences in computed conformational preferences, which further suggests that macrolide conformation plays an important role in the low nanomolar biological activity of neopeltolide. Unfortunately, the lack of alternative solution conformations offers us limited opportunity for analogue design, especially with the current state of SAR studies covering virtually every possible modification to both the macrolide and side chain. As such, future studies on neopeltolide would be better served by focusing on neopeltolide’s mode of action.

Though Kozmin and co-workers established that neopeltolide and leucascandrolide A inhibit the function of cytochrome bc1, subsequent literature raises doubts as to whether this would account for its anticancer activity. As covered previously, the role of cytochrome bc1 in cancer cell proliferation has not been conclusively linked, and known inhibitors of the complex are generally antifungal agents rather than anticancer. A recent mode-of-action study by Fuwa using 8,9-
dehydroneopeltolide on HL-60 human leukemia cells found that the compound appears to inhibit the electron transport chain as expected, inducing apoptosis through the release of cytochrome c from the mitochondria. However, this required the simultaneous use of a glycolysis inhibitor to sensitize the cells, and the pathway they observed for triggering cytotoxicity was different from that of the known Q_{10}-site inhibitor antimycin A.\textsuperscript{334}

Furthermore, work by Singh and co-workers found that the activity of cytochrome $b_{c_1}$ in metastatic and aggressive breast cancer cell lines is already reduced significantly (including the MCF7 line used in several neopeltolide assays), and concluded that impairment of this complex actually contributes to the development of breast cancer.\textsuperscript{335}

These puzzling findings and contradictions suggest that neopeltolide’s mode of action is likely not the mere inhibition of cytochrome $b_{c_1}$, warranting a closer look at other potential targets or vectors.

6.1.1 Neopeltolide as a Fluorescent Probe

Our conformational studies on C2 substitution indicate that modifications to this site are unlikely to alter macrolide conformation to any significant degree, which means that the same methodology can be used to attach substituents for alternative purposes. With neopeltolide’s mode of action remaining an open question, it would be greatly beneficial to have some way of tracking the compound as it works its way through a cell. One way to do this is via fluorescent tagging, turning the compound into a probe. La Clair and co-workers found success with a central dye based on coumarin \textbf{6.1}, which they found to be readily taken up into cells while possessing no intrinsic biological activity or subcellular localization of its own.\textsuperscript{336} Additionally, attachment of this label had minimal effect on their sample natural products, which included a number of macrocyclic
polyketides. Their methodology includes a means of converting 6.1 to a variety of functionalized labels through modification of the ester. While none of them provides an electrophilic handle, their alkene 6.2 could be converted to a terminal iodide 6.3 using methodology developed by Renaud, as shown in Scheme 6.1. From here, attaching the probe to the neopeltolide macrolide should be possible with the same method we used to generate 2-methyl neopeltolide, which would likely yield the same $R$ configuration as detailed earlier. La Clair and co-workers found that their coumarin labeling allowed for the studying of membrane association and for screening possible secondary systems via the labeling of amphotericin B and avermectin, respectively, and could provide similar insights for a labeled neopeltolide.

![Scheme 6.1](image)

**Scheme 6.1 – Conversion and attachment of coumarin-based dye 6.1.** a) LiOH, THF, H$_2$O; b) pent-4-en-1-ol, EDAC, DMF; c) CatBH, MeCONMe$_2$ (cat.); d) $p$-TolSO$_2$I, $t$-BuON=NO= $t$-Bu; e) NaHMDS (5 eq), then 6.3.

### 6.1.2 Re-Evaluation of Cytochrome $bc_1$

One of the more glaring omissions in Kozmin’s mode of action study is its lack of controls. As mentioned previously, the shared side chain 3.4 possesses antifungal activity
but not anticancer activity, which would make it a useful marker for distinguishing potentially different modes of action. As cytochrome \(bc_1\) is a well-known antifungal target, its inhibition by neopeltolide may be responsible for the observed antifungal activity but not necessarily its anticancer activity. Additionally, \textbf{3.4} bears a decent resemblance to coenzyme Q10 and some of the known \(Q_o\)-site inhibitors such as myxothiazol and famoxadone. As such, it’s possible that direct treatment of cytochrome \(bc_1\) as described by Kozmin with \textbf{3.4} could give similar results to what was observed for neopeltolide and leucascandrolide A, which would indicate that simple inhibition of this complex is insufficient for anticancer activity.

![Figure 6.1 – Comparison of \textbf{3.4} to known \(Q_o\)-site inhibitors famoxadone and myxothiazol.](image)

6.2 Zampanolide

Our conformational study of zampanolide and dactylolide’s macrolide core used a combination of high-field NMR experiments and molecular modeling to, which found
that it possesses three interconverting conformational families while in solution. One of these families bore a striking resemblance to the bound conformation of zampanolide, with only the only difference being the orientation of the C9-C10 enone. Given our suspicion that the macrolide seems to be serving little purpose other than to provide space between the M-loop interactions of the side chain and the covalent bond formed with the enone, we designed a simplified analogue that should be easier to synthesize yet hopefully retain biological activity. According to our theory, additional conformationally-related simplifications should also be possible without adversely affecting biological activity, namely removing the northern diene and the C5 methyl group.

However, while SAR studies to date have focused on modifications to the side chain and macrolide, the importance of the covalent binding of zampanolide and dactylolide has been virtually unexplored. Recently, our lab had opportunity to test zampanolide against multiple cancer cell lines and several unique MSA-resistant mutants. These tests found that zampanolide displays potent cytotoxicity against P-glycoprotein expressing ovarian cancer cells and both epothilone and peloruside-resistant cell lines.\(^{301}\) At the same time, my colleague Matt Wilson took the opportunity to test a precursor compound of dactylolide (6.4) against PC-3 cells that would be unable to undergo Michael addition to form the C9 covalent bond. Unsurprisingly, this compound displayed significantly reduced cytotoxicity when compared to a known simplified dactylolide analogue 6.5,\(^{292}\) which indicates that covalent binding plays a significant role in at least dactylolide’s potent biological activity.
6.2.1 Non-Covalent Zampanolide Analogue

As mentioned previously, Field and co-workers found that both zampanolide and dactylolide are capable of rapidly displacing Flutax-2 from the taxane binding pocket on β-tubulin. However, dactylolide remains in exchange with the medium and takes significantly longer to covalently bind with H229, while zampanolide reacts almost immediately. This suggests that the hemiaminal side chain plays a role in binding kinetics, possibly solidifying itself in the pocket well enough via non-covalent interactions to facilitate easy covalent binding. Inspection of the co-crystal structure reveals that the side chain appears to be inserting itself into a hydrophobic pocket within the M-loop. As such, it would be interesting to see if zampanolide binds well enough to the taxane site to retain biological activity without covalently binding.

Fortunately, modification of zampanolide to remove covalent binding ability can be achieved without significant modification to our lab’s synthetic route. The C7 carbonyl is installed as part of the northern diene fragment 4.3, which is deprotected and oxidized following its coupling to the tetrahydropyran fragment. Since this oxidation
normally coincides with the formation of the C20 aldehyde, we would instead convert the precursor alcohol 6.6 to the methyl ester 6.7 using Meerwein’s reagent. This in turn can be converted to the carboxylic acid 6.8, which can subsequently be coupled to the tetrahydropyran fragment using Jennings’s methodology and finished to the zampanolide analogue 6.9. If the non-covalent analogue proves to possess biological activity, we could also incorporate the modified fragment into our conformationally-simplified zampanolide analogue. This analogue would allow us to directly assess the effects of conformation on zampanolide’s non-covalent binding to the taxane site, and determine whether or not the rigidifying $A^1,3$-strain conferred by the C17 methyl plays a role in biological activity.

Scheme 6.2 – Route to non-covalent analogue 6.9 and its differences from the standard zampanolide route.
CHAPTER 7:
SUPPORTING INFORMATION

7.1 General Methods

Unless otherwise noted, all materials were used as received from a commercial supplier without further purification. All anhydrous reactions were performed using oven-dried or flame-dried glassware, cooled under vacuum and purged with argon gas. Dichloromethane (DCM), tetrahydrofuran (THF), toluene and diethyl ether (Et₂O) were filtered through activated alumina under nitrogen. Triethylamine (Et₃N) was distilled over CaH₂ and stored over KOH pellets. 4 Å molecular sieves were oven-dried overnight and cooled under high vacuum prior to use. All reactions were monitored by Silicycle thin layer chromatography (TLC) plates (Extra Hard Layer, 60 Å, glass back) and analyzed with 254 nm UV light and/or anisaldehyde treatment. Silica gel for column chromatography was purchased from Silicycle (SiliaFlash® P60, 230 – 400 mesh).

Unless otherwise noted, all ¹H and ¹³C NMR spectra were recorded in CDCl₃ using either a Varian Inova 500 spectrometer operating at 499.86 MHz for ¹H and 125.69 MHz for ¹³C, or Varian VNMRS 600 operating 599.87 MHz for ¹H and 150.84 MHz for ¹³C. Chemical shifts (δ) were reported in ppm relative to the residual CHCl₃ as an internal reference (¹H: 7.26 ppm, ¹³C: 77.23 ppm). Coupling constants (J) were reported in Hertz (Hz). Peak multiplicity is indicated as follows: s (singlet), d (doublet), t (triplet), q (quartet), p (pentet), x (septet), h (heptet), b (broad) and m (multiplet). Mass spectra
(FAB) were obtained at the Department of Chemistry and Biochemistry, University of Notre Dame using either a JEOL AX505HA or JEOL JMS-GCmate mass spectrometer.

**7.2 Experimental Methods**

**TBS-Protected Macrolide**

![TBS-Protected Macrolide](image)

To a solution of macrolide 3.57 (40 mg, 0.122 mmol) in CH$_2$Cl$_2$ (3 mL) at -78°C was added 2,6-lutidine (28.4 µL, 0.244 mmol) and TBSOTf (33 µL, 0.183 mmol) sequentially. The reaction mixture was stirred for 2 h at -78°C and then quenched with saturated aqueous NH$_4$Cl. After warming to room temperature, the layers were separated and the aqueous layer further extracted with CH$_2$Cl$_2$ (4 x 10 mL). The combined organic layers were dried with MgSO$_4$, filtered, and concentrated *in vacuo*. Purification by column chromatography (95:5 Hexanes:EtOAc) gave the TBS-protected macrolide (50 mg, 93% yield) as a colorless oil. $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 5.13 (td, $J = 9.4, 4.8$ Hz, 1H), 4.17 – 4.09 (m, 2H), 3.69 – 3.61 (m, 1H), 3.61 – 3.53 (m, 1H), 3.31 (s, 3H), 2.55 (dd, $J = 14.2, 4.6$ Hz, 1H), 2.50 (dd, $J = 14.2, 10.8$ Hz, 1H), 1.95 – 1.86 (m, 1H), 1.79 – 1.69 (m, 1H), 1.60 – 1.55 (m, 1H), 1.53 (d, $J = 2.6$ Hz, 1H), 1.52 (d, $J = 2.2$ Hz, 2H), 1.43 – 1.24 (m, 8H), 1.18 (dd, $J = 14.9, 2.1$ Hz, 1H), 1.12 (ddd, $J = 12.9, 10.7, 2.2$ Hz, 1H), 0.97 (d, $J = 6.8$ Hz, 3H), 0.91 (t, $J = 5.4$ Hz, 3H), 0.90 – 0.89 (m, 9H). $^{13}$C NMR (126 MHz, CDCl$_3$) $\delta$ 171.57, 75.97, 74.83, 73.52, 69.36, 65.54, 56.45, 44.46, 42.91, 42.54,
218

40.33, 40.30, 39.55, 37.18, 31.30, 30.57, 26.09, 25.75, 19.25, 18.35, 14.16. HRMS-FAB: 
(M+H)$^+$ = 443.3187 calculated for C$_{24}$H$_{47}$O$_5$Si, experimental: 443.3191.

TBS-Protected (R)-2-Methyl Macrolide

To a solution of TBS-protected macrolide (11.5 mg, 0.026 mmol) in THF (0.5 mL) at -78°C was added 65 µL of NaHMDS (2.0 M in THF, 0.13 mmol). The solution was stirred at -78°C for 2 h, then MeI (10 µL, 0.156 mmol) was added and the mixture warmed to 0°C. After 1 h of further stirring, full conversion was observed and the reaction was quenched with 1 mL of H$_2$O. 3 mL each of saturated aqueous NaHCO$_3$ and EtOAc were added and the organic layer was separated. The aqueous layer was further extracted with EtOAc (2 x 2 mL), and the combined organic layers were washed with brine (3 mL), then dried with MgSO$_4$, filtered, and concentrated in vacuo. Purification by column chromatography (95:5 Hexanes:EtOAc) gave the methylated macrolide (10 mg, 84% yield) as a colorless oil. $^1$H NMR (600 MHz, CDCl$_3$) δ 4.98 (s, 1H), 4.20 – 4.12 (m, 1H), 3.69 (ddd, $J$ = 11.6, 9.7, 2.0 Hz, 1H), 3.61 (dt, $J$ = 9.8, 6.8 Hz, 1H), 3.54 (dddd, $J$ = 10.8, 9.4, 2.5, 1.3 Hz, 1H), 3.32 (s, 3H), 2.41 (dq, $J$ = 9.8, 7.0 Hz, 1H), 2.09 – 1.98 (m, 1H), 1.89 – 1.78 (m, 1H), 1.73 – 1.65 (m, 1H), 1.60 – 1.47 (m, 3H), 1.41 (dd, $J$ = 7.3, 2.8 Hz, 2H), 1.39 – 1.37 (m, 1H), 1.37 – 1.22 (m, 7H), 1.18 – 1.12 (m, 1H), 1.12 – 1.06 (m, 4H), 0.96 (d, $J$ = 6.9 Hz, 3H), 0.92 (t, 3H), 0.90 (m, 9H). $^{13}$C NMR (151 MHz, CDCl$_3$) δ
To a solution of macrolide (9.0 mg, 0.0197 mmol) in THF (0.7 mL) was added TBAF (1.0 M in THF, 98 µL, 0.098 mmol). After 2 h, additional TBAF (98 µL) was added, and this procedure was repeated to a total reaction time of 6 h. The reaction mixture was quenched with 2 mL saturated aqueous NH₄Cl, then extracted with EtOAc (3 x 3 mL). The organic layers were combined, dried with MgSO₄, filtered, and concentrated in vacuo. Purification by column chromatography (90:10 to 70:30 Hexanes:EtOAc) gave the macrolide alcohol 3.58 (5.2 mg, 77%) as a colorless oil. 

$^1$H NMR (600 MHz, CDCl₃) δ 5.06 (s, 1H), 4.29 – 4.23 (m, 1H), 3.76 (ddd, $J = 11.8, 10.0, 2.1$ Hz, 1H), 3.64 (td, $J = 9.8, 3.1$ Hz, 1H), 3.58 – 3.50 (m, 1H), 3.32 (s, 3H), 2.45 (dq, $J = 9.9, 7.0$ Hz, 1H), 2.01 – 1.91 (m, 1H), 1.84 – 1.71 (m, 2H), 1.61 – 1.47 (m, 5H), 1.42 (m, 3H), 1.38 – 1.28 (m, 3H), 1.22 – 1.17 (m, 1H), 1.15 – 1.04 (m, 4H), 0.97 (d, $J = 6.8$ Hz, 3H), 0.91 (t, $J = 7.4$ Hz, 3H). 

$^{13}$C NMR (126 MHz, CDCl₃) δ 174.95, 75.94, 75.18, 74.81, 73.44, 65.04, 56.35, 48.26, 44.24, 42.23, 40.17, 39.71, 37.04, 36.75, 31.11, 29.91,
25.54, 19.23, 14.10. HRMS-FAB: \((\text{M}+\text{Na})^+ = 365.2298\) calculated for \(\text{C}_{19}\text{H}_{34}\text{O}_5\), experimental: 365.2295. \([\alpha]_{\text{D}}^{20} = +20.5^\circ\) (c = 0.50 in CHCl\(_3\)).

(R)-2-Methyl Neopeltolide (3.68)

![Chemical structure of R-2-Methyl Neopeltolide](image)

To a solution of 3.58 (6.0 mg, 18 \(\mu\)mol) and bis(2,2,2-trifluoroethyl)phosphonoacetic acid (11 mg, 36 \(\mu\)mol) in CH\(_2\)Cl\(_2\) (2.5 mL) was added HOBT·H\(_2\)O (1.0 mg, 7 \(\mu\)mol) and EDCI·HCl (33.6 mg, 180 \(\mu\)mol). After 20 min, the mixture was filtered through a short plug of silica gel, eluting with 50 mL EtOAc. The eluant was concentrated and the residue was used immediately without further purification.

To a cooled (-78°C) solution of 18-crown-6 ether (24 mg, 90 \(\mu\)mol) and the unpurified phosphonate in THF (0.5 mL) was added potassium bis(trimethylsilyl)amide (0.5M in toluene, 45 \(\mu\)L, 21.6 \(\mu\)mol). After 1 h the reaction mixture was cooled to -85°C and a solution of aldehyde 3.65 (9 mg, 36 \(\mu\)mol) in THF (0.5 mL) was added dropwise by cannula. The reaction mixture was stirred for 3 h at -85°C and then quenched by the addition of saturated aqueous NH\(_4\)Cl (5 mL). The aqueous layer was extracted with Et\(_2\)O (4 x 5 mL), and the combined organic layers were dried over MgSO\(_4\), filtered, and concentrated \textit{in vacuo}. Preparative TLC (50:50 Hexanes:EtOAc) afforded 6.0 mg (57% yield over 2 steps) of 3.68, judged to be a 4:1 mixture of Z:E isomers by \(^1\)H NMR.
analysis. $^1$H NMR (500 MHz, CDCl$_3$) δ 7.39 (d, $J = 1.2$ Hz, 1H), 6.40 – 6.22 (m, 2H), 6.17 – 6.03 (m, 1H), 5.88 (dt, $J = 11.5$, 1.7 Hz, 1H), 5.53 (s, 1H), 5.27 – 5.16 (m, 1H), 4.97 (s, 1H), 4.32 (t, $J = 6.5$ Hz, 2H), 3.77 – 3.58 (m, 4H), 3.58 – 3.42 (m, 3H), 3.32 (s, 3H), 3.04 (qd, $J = 7.4$, 1.8 Hz, 2H), 2.72 (td, $J = 7.5$, 1.2 Hz, 2H), 2.45 (dq, $J = 9.7$, 7.0 Hz, 1H), 2.02 (t, $J = 13.0$ Hz, 1H), 1.92 (d, $J = 13.6$ Hz, 1H), 1.84 (h, $J = 8.8$ Hz, 1H), 1.70 – 1.63 (m, 1H), 1.55 – 1.47 (m, 3H), 1.46 – 1.29 (m, 7H), 1.29 – 1.14 (m, 5H), 1.14 – 1.04 (m, 4H), 0.97 (t, $J = 7.3$ Hz, 3H), 0.92 (t, $J = 7.3$ Hz, 3H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 174.88, 165.62, 149.34, 136.48, 134.07, 120.93, 116.89, 75.92, 75.31, 67.62, 66.05, 56.35, 48.13, 44.08, 42.13, 39.86, 39.55, 36.92, 36.54, 33.90, 30.67, 29.91, 27.87, 25.91, 25.35, 19.33, 15.48, 14.08, 13.96. HRMS-FAB: (M+Na) $^+$ = 627.3252 calculated for C$_{32}$H$_{48}$N$_2$O$_9$, experimental: 627.3282. [$\alpha$]$^20_D$ = +22.0° (c = 0.20 in CHCl$_3$).
### 7.3 NMR Data

**TABLE 7.1**

NEOPELTOLIDE MACROLIDE $^1$H NMR COUPLING CONSTANTS

<table>
<thead>
<tr>
<th>Proton</th>
<th>Chemical Shift (ppm) and Splitting</th>
<th>$J$ (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2proR</td>
<td>2.56, dd</td>
<td>14.4, 4.2</td>
</tr>
<tr>
<td>2proS</td>
<td>2.32, dd</td>
<td>15.0, 11.4</td>
</tr>
<tr>
<td>3</td>
<td>4.2, dddd</td>
<td>11.4, 11.4, 4.8, 2.4</td>
</tr>
<tr>
<td>4proR</td>
<td>1.67, m</td>
<td>–</td>
</tr>
<tr>
<td>4proS</td>
<td>1.5, m</td>
<td>–</td>
</tr>
<tr>
<td>5</td>
<td>4.25, b</td>
<td>–</td>
</tr>
<tr>
<td>6proR</td>
<td>1.5, m</td>
<td>–</td>
</tr>
<tr>
<td>6proS</td>
<td>1.67, m</td>
<td>–</td>
</tr>
<tr>
<td>7</td>
<td>3.69, ddd</td>
<td>11.4, 9.0, 2.4</td>
</tr>
<tr>
<td>8proR</td>
<td>1.21, dd</td>
<td>15.6, 1.8</td>
</tr>
<tr>
<td>8proS</td>
<td>1.39, ddd</td>
<td>15.0, 9.0, 5.4</td>
</tr>
<tr>
<td>9</td>
<td>1.5, m</td>
<td>–</td>
</tr>
<tr>
<td>10proR</td>
<td>1.58, ddd</td>
<td>13.2, 11.4, 2.4</td>
</tr>
<tr>
<td>10proS</td>
<td>1.12, ddd</td>
<td>13.2, 10.8, 2.4</td>
</tr>
<tr>
<td>11</td>
<td>3.6, dddd</td>
<td>10.8, 9.6, 2.4, 1.2</td>
</tr>
<tr>
<td>12proR</td>
<td>1.35, m</td>
<td>–</td>
</tr>
<tr>
<td>12proS</td>
<td>1.83, ddd</td>
<td>15.0, 10.8, 1.8</td>
</tr>
<tr>
<td>13</td>
<td>5.2, dddd</td>
<td>9.6, 9.6, 4.8, 0.6</td>
</tr>
</tbody>
</table>

*NOTE: All coupling constants taken in CDCl$_3$. Original data is from Kartika et al, *Org. Lett.*, 2008, 10, 5047. Specific coupling constants used for DISCON calculations are found in Figure 3.25.*
TABLE 7.2
DACTYLOLIDE ROESY-DERIVED INTERPROTON DISTANCES USED IN
DISCON CALCULATIONS

<table>
<thead>
<tr>
<th>H1</th>
<th>H2</th>
<th>Peak Volume</th>
<th>Interproton distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10proR</td>
<td>10proS</td>
<td>1.00</td>
<td>1.78</td>
</tr>
<tr>
<td>16</td>
<td>14proR</td>
<td>0.11</td>
<td>2.65</td>
</tr>
<tr>
<td>18proR</td>
<td>16</td>
<td>0.27</td>
<td>2.32</td>
</tr>
<tr>
<td>10proR</td>
<td>8</td>
<td>0.06</td>
<td>2.94</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>0.07</td>
<td>2.81</td>
</tr>
<tr>
<td>9</td>
<td>11</td>
<td>0.09</td>
<td>2.79</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>0.14</td>
<td>2.65</td>
</tr>
<tr>
<td>3</td>
<td>6proS</td>
<td>0.27</td>
<td>2.30</td>
</tr>
<tr>
<td>3</td>
<td>6proR</td>
<td>0.13</td>
<td>2.56</td>
</tr>
<tr>
<td>6proS</td>
<td>9</td>
<td>0.12</td>
<td>2.69</td>
</tr>
<tr>
<td>17Me</td>
<td>15</td>
<td>0.29</td>
<td>2.22</td>
</tr>
<tr>
<td>17Me</td>
<td>19</td>
<td>0.18</td>
<td>2.42</td>
</tr>
</tbody>
</table>
TABLE 7.3
DACTYLOLIDE $^1$H NMR COUPLING CONSTANTS USED IN DISCON CALCULATIONS

<table>
<thead>
<tr>
<th>H1</th>
<th>H2</th>
<th>J (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10proS</td>
<td>9</td>
<td>7.5</td>
</tr>
<tr>
<td>10proS</td>
<td>11</td>
<td>2.5</td>
</tr>
<tr>
<td>19</td>
<td>18proR</td>
<td>11.2</td>
</tr>
<tr>
<td>19</td>
<td>18proS</td>
<td>2.7</td>
</tr>
<tr>
<td>10proR</td>
<td>11</td>
<td>10.0</td>
</tr>
<tr>
<td>10proR</td>
<td>9</td>
<td>6.9</td>
</tr>
</tbody>
</table>

APPENDIX A:
ELECTRONIC SUPPLEMENT

Available with this dissertation is an electronic supplement containing various input and output files related to the calculations discussed. The directories are organized according to the Chapters and Sections in which the relevant calculations are first covered. For Monte Carlo conformational searches, the necessary files to reproduce calculations, such as input geometries and command files, are included. For MD calculations, necessary files to reproduce calculations, such as model system files and input coordinate files, are included. For Chapter 5, Mercury CSD files have been included, along with Excel spreadsheets where data was combined to reflect multiple searches. DISCON input and output files have been included, along with a copy of the DISCON software in case the source download is unavailable (note: all DISCON file directories must be copied to “C:\DISCON\[Folder Name]” in order to be read).
APPENDIX B:

NMR SPECTRA
Pulse Sequence: ROESY

Solvent: desco
Temp: 22.0 °C / 253.1 K
Operator: mwllso16
File: Dactylolide2DROESY800ms

NMR-500 "nmr500"

Relax. delay 4.000 sec
Mixing 6.000 sec
Acq. time 0.177 sec
Width 5787.0 Hz
2D Width 5787.0 Hz
32 repetitions
2 x 128 increments

OBSERVE: H1, 599.8756950 MHz
DATA PROCESSING
Gauss apodization 0.002 sec
F1 DATA PROCESSING
Gauss apodization 0.001 sec
FT size 256 x 256
Total time 11 hr. 25 min. 53 sec

(-)-dactylolide 4.14
File: afs/nd.edu/user28/mw15016/Private/NMR/Dactylolide2DROESY60ms-B.fld

Pulse Sequence: ROESY
Solvent: deo
Temp: 22.0°C / 295.1 K
Operator: mw15016
File: Dactylolide2DROESY688ms-B

Relax. delay 4.800 sec
Mixing 0.800 sec
Acq. time 0.177 sec
Width 5787.0 Hz
2D Width 5787.0 Hz
32 repetitions
2 x 102 increments
OBSERVE: H1, 559.8757665 MHz
DATA PROCESSING
Gauss apodization 0.083 sec
FI DATA PROCESSING
Gauss apodization 0.141 sec
FT size 2048 x 2048
Total time 21 hr, 25 min, 53 sec

(-)-dactylolide 4.14
File: afs/md.edu/user25/mwilsol6/Private/NMR/Dactylolide2DROESYB08ms-B.fid

Pulse Sequence: ROESY
Solvant: deca
Temp: 22.0 °C / 295.1 K
Operator: mwilsol6
File: Dactylolide2DROESYB08ms-B
Vnmrs-600 "nmr600"

Relax. delay 4.000 sec
Mixing 0.800 sec
Acq. time 0.137 sec
Width 5787.9 Hz
2D Width 5787.9 Hz
32 repetitions
2 x 128 increments
OBSERVE H1, 598.875769 MHz
DATA PROCESSING
Gauss apodization 0.002 sec
F1 DATA PROCESSING
Gauss apodization 0.041 sec
FT size 2048 x 2048
Total time 11 hr, 25 min, 53 sec

Std proton

(-)-dactylolide 4.14
File: afs.nd.edu/user23/mwilso16/Private/NMR/Dactyloolide2DROESY800ms-B.tif
Pulse Sequence: ROESY
Solvent: d2o
Temp. 22.0 C / 295.1 K
Operator: mwilso16
File: Dactyloolide2DROESY800ms-B

VNMRS-600 "mer600"

Relax. delay 4.000 sec
Mixing 0.800 sec
Acq. time 0.077 sec
Width 5787.6 Hz
ZB Width 5787.6 Hz
32 repetitions
2 x 128 increments

OBSERVE H1, 599.8757069 MHz
DATA PROCESSING
Gauss apodization 0.002 sec f1 DATA PROCESSING
Gauss apodization 0.001 sec f2 DATA PROCESSING
FT size 2048 x 2048
Total time 11 hr, 25 min, 53 sec

(-)-dactyloolide 4.14

H2-H4

H3-H9
BIBLIOGRAPHY

(1) Baeyer, A. Berichte der Deutschen Chemischen Gesellschaft 1885, 18, 2269.

(2) Sachse, H. Berichte der Deutschen Chemischen Gesellschaft 1890, 23, 1363.

(3) Russell, C. A. In van't Hoff-Le Bel Centennial; AMERICAN CHEMICAL SOCIETY: 1975; Vol. 12, p 159.

(4) Mohr, E. Journal für Praktische Chemie 1918, 98, 315.

(5) Mohr, E. Journal für Praktische Chemie 1921, 103, 316.


(7) Hückel, W. Justus Liebigs Annalen der Chemie 1925, 441, 1.

(8) Hassel, O. Bergvesen og Metallurgi 1943, 3, 32.

(9) Bastiansen, O.; Hassel, O. Bergvesen og Metallurgi 1946, 8, 96.


(11) Barton, D. Experientia 1950, 6, 316.


(16) For information on the bryostatins in clinical trials, see: http://clinicaltrials.gov.


(64) Cossy, J. *Comptes Rendus Chimie* **2008**, *11*, 1477.


(74) Smith, A. B.; Freeze, B. S. *Tetrahedron* **2008**, *64*, 261.


(146) DISCON software and documentation is available for download at http://discon.sourceforge.net.


(171) Lang, W.; Caldwell, G. W.; Li, J.; Leo, G. C.; Jones, W. J.; Masucci, J. A. *Drug Metabolism and Disposition* 2007, 35, 21.

(172) Guo, W.; Reigan, P.; Siegel, D.; Ross, D. *Drug Metabolism and Disposition* 2008, 36, 2050.


(208) Nicholson, C. P. A complete program toward polyketide natural products: Methodology, synthesis, and conformational analysis; University of Notre Dame, 2011.


(222) Youngsaye, W. W.-M., Boston University, 2010.


(232) Vintonyak, V. V.; Maier, M. E. Org. Lett. 2008, 10, 1239.


(234) Tu, W.; Floreancig, P. E. Angewandte Chemie 2009, 121, 4637.


(301) Wilson, M. R., University of Notre Dame, **2014**.


