Design, Synthesis and Evaluation of Cholera Toxin Inhibitors and α-Helix Mimetics of Dormancy Survival Regulator

Guangtao Zhang

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

Doctor of Philosophy

University of Washington

2006

Program Authorized to Offer Degree:
Chemistry

University of Washington
Graduate school

This is to certify that I have examined this copy of a doctoral dissertation by

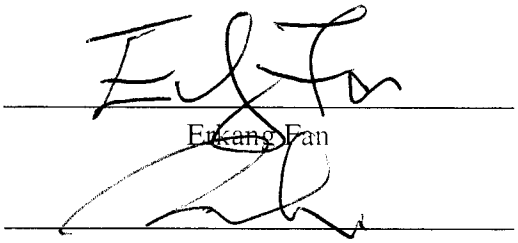
Guangtao Zhang

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of the Supervisory Committee:

Erkang Fan

Reading Committee:



Erkang Fan

Tomikazu Sasaki

Xiaosong Li

Date: 12/8/2006

In presenting this dissertation in partial fulfillment of the requirements for the doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, 1-800-521-0600, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature Zhang Guangtao

Date 12/8/2006

University of Washington

Abstract

Design, Synthesis and Evaluation of Cholera Toxin Inhibitors and α-Helix Mimetics of Dormancy Survival Regulator

Guangtao Zhang

Chair of the Supervisory Committee:
Professor Erkang Fan
Department of Biochemistry

Infectious diseases cause millions of deaths each year. The top single disease killers include tuberculosis (TB) and diarrheal diseases. The diarrheal diseases caused by cholera toxin (CT) and *Escherichia coli* heat-labile enterotoxin (LT) constitute a large part of the diarrhea-related deaths worldwide each year. While there have been some efficacious drugs against *Vibrio cholerae* in the market, their resistance is developing very quickly. Our group is engaged in developing synthetic molecules that could bind to CT with a high affinity to block the enzymatic reaction catalyzed by CT. We report here the design, synthesis and biological evaluation of bisubstrate analogues targeting the active site of CT. The most potent inhibitor demonstrated an IC50 of around 40 μM, which is 350-fold more potent than the natural substrate NAD. This compound could potentially serve as a lead for the discovery of therapeutic reagent to treat cholera.

The other project is the exploration of the use of α-helix mimetics for probing the protein-protein interactions of dormancy survival regulator (DosR). The main function of DosR is to mediate the transition of *Mycobacteria tuberculosis* into dormancy and may contribute to latency. Disruption of the dimerization of DosR could keep *M. tuberculosis* in an active state, so that the existing drugs could treat them effectively. We employed

both the small molecule approach and peptide approach to design potent α-helix mimetics and so far the most potent compound showed an IC50 of around 40 μM. The long linear peptide inhibitors synthesized in this study demonstrated an IC50 of around 100 μM. This opens up the possibility of systematic investigation of the DosR interface by incorporating different side chains. This strategy could be extended to study other interfaces with protein-protein interaction.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYNTHETIC SCHEMES

LIST OF TABLES

LIST OF ABBREVIATIONS

2D	Two dimensional
3D	Three dimensional
Å	Angstrom
Abz	Aminobenzoic acid
ADP	Adenosine diphosphate
ADPR	ADP-ribosylation reaction
ADPR	ADP-ribosylation reaction
ADPRTs	ADP-ribosylating toxins
Aib	2-Aminoisobutyric acid
ARF	ADP-ribosylation factor
BSA	Bovine serum albumin
CAMM	Computer-assisted molecular modeling
cAMP	Cyclic adenosine monophosphate
CD	Circular dichroism
CSD	Chemical shift deviation
CT	Cholera toxin
CTA	Cholera toxin A subunit
Dap	L-α,β -diaminopropionic acid
DEABAG	Diethylamino(benzylidine-amino)guanidine
DIPEA	Diisopropylethylamine
DLS	Dynamic light scattering

DMA	N,N-dimethyl acetamide
DMF	N,N-dimethylformamide
DMSO	Methyl sulfoxide
DosR	Dormancy survival regulator
$DosR^c$	Dormancy survival regulator c-terminal domain
dppf	1,1'-Bis(diphenylphosphino)ferrocene
DSS	2,2-dimethyl-2-silapentane- 5-sulfonate sodium salt
E. coli	*Escherichia coli*
EDC	1-[3-(Dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride
Eps	Extracellular protein secretion
ER	Endoplasmic reticulum
ETEC	Enterotoxigenic *Escherichia coli*
EtOAc	Ethyl acetate
Fmoc	9-Fluorenylmethoxycarbonyl
FT	Fourier transformation
GM1	Ganglioside M1
$G_{s\alpha}$	Alpha subunit of the stimulatory G protein
Gsp	General secretion pathway
GTP	Guanosine triphosphate
HIV	Human immunodeficiency virus
HOBt	Hydroxybenzotriazole
HPLC	High-performance liquid chromatography
hspX	Hypoxic response gene *acr*

HTC	High-throughput crystallography
HTC	high-throughput screening
IC50	Inhibitory concentration (corresponding to 50% inhibition)
LT	Heat-labile exterotoxin
LTB	Latent TB
MBHA	Methylbenzhydrylamine
MDM2	Murine double-minute 2
MS	Mass Spectrum
MTB	Mycobacteria tuberculosis
MTIP	MyoA-tail Interacting Protein
Mtr	4-Methoxy-2,3,6-trimethylbenzene-sulfonyl
NAD	Nicotinamide adenine dinucleotide
NMR	Nuclear magnetic resonance
ORT	Oral rehydration therapy
Pbf	2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl
PDI	Protein disulfide isomerase
PPD	Purified protein derivative
PT	Pertussis toxin
PyBOP	benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
ROESY	Rotating frame Overhause Effect SpectroscopY
SAR	Structure activity relationship
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SHT	Shiga toxin
SLT	Shiga-like toxins
TB	Tuberculosis
TBE	Tris/Borate/EDTA
TFA	Trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
THF	Tetrahydrofuran
TIS	Triisopropylsilane
TM	Trade mark
TOCSY	TOtal Correlated SpectroscopY
UV	Ultraviolet
WHO	World Health Organization

ACKNOWLEDGEMENT

I would like to thank my advisor Professor Erkang Fan for his guidance and advices in both research and life during the years I work in this laboratory. I would also like to thank Professor Wim Hol and Professor Christophe Verlinde for stimulating discussions about structure-based drug design and computer modeling. I would also like to thank my wonderful collaborators Goragot Wisedchaisri, Claire O'Neal, David Roberts, David Sherman.

I would like to show my sincere gratitude to the following friends and colleagues for their helps and supports during my graduate study and Ph.D. research at the University of Washington: Tyan Carter, Yu Lu, Chu Wang, Katherine Olsen, Claire O'Neal, Rajan Paranji, Jason Pickens, Goragot Wisedchaisri, Yadong Yin, Zhongsheng Zhang.

Finally, I would like to thank my committee members: Drs. Erkang Fan, Daniel Gamelin, Wim Hol, Xiaosong Li, Tomikazu Sasaki for their invaluable suggestions and comments.

DEDICATION

I wish to dedicate this thesis to my parents for their unconditional support and love, and to Uncle Chen and Aunt Chen for their love and care. I would also like to dedicate this thesis to my girlfriend and my best friend, Yeon Jeong Lee, for her love and support during this period of study.

Chapter 1

Background and Introduction

Structure-based Drug Design

For decades, drug discovery has been closely associated with random and accidental findings. The general strategy is to screen a wide array of compounds from either commercial databases or analogue development programs against the target. The top hit is identified and is called a "lead compound". Through reiterative cycles, lead compound is systematic modified using medicinal chemistry methods and subsequent testing. The vast majority of drugs currently in the market are developed this way. However, the drug development cycle is very long —— it took up to 6-12 years to successfully put a drug through to the market.[1] This expensive and lengthy process of searching and developing for therapeutic agents could not keep up with the demand. A turning point from serendipity to rational drug design was at hand. In the meantime, recent progress in cloning, expressing and purifying proteins has been enormous, which makes it possible to obtain proteins in large quantities.[2] Site-directed mutagenesis allows us to make specific changes to the enzyme, hormone and other cofactors.[3-9] Advances in structural techniques open up the possibilities to determine protein structures at atomic level by either high-resolution X-ray crystallography or nuclear magnetic resonance.[10-11] These profound technical advancements lead to a protein structure guided drug development approach —— Structure-based drug design.[1,2,12] The primary goal of structure-based drug design is to optimize ligand potency and affinity, yet it is not limited to that. Drug design encompasses optimization of many other properties such as

dissolution, absorption, metabolic stability, plasma protein binding, distribution, elimination, toxicological profile and pharmacokinetics and pharmacodynamics properties.[13] Within the scope of this study, it would be more accurate to term this approach as "structure-based inhibitor design".

A typical structure-based drug design cycle starts with a three-dimensional structure of any form of the receptor produced by X-ray crystallography (HTC).[1,2,12] Technical advances have made the dream of high-throughput crystallography come true.[14] In particular, synchrotron beam lines have become more intense and focused, which enabled rapid data collection. The throughput capacities have been increased up to 140,000 experiments per day.[14] An example in case is the development of inhibitors against DPP4 in Takeda San Diego — by employing high-throughput crystallography, [15] they were able to obtain thirteen crystal structures of their target protein complex within a month. Over a span of three months, they got thirty structures solved which facilitated SYR-322 to be in phase II clinical trial within one year.[16] Technological innovations have allowed previously unthought of experiments to be carried out and X-ray crystallography almost turns into a screening tool. More importantly, it affords a large number of structures of disease-related proteins. NMR methods have also produced many new structures, mostly small and rigid proteins. The major advantage is that proteins need not be crystallized and flexible linker regions associated with globular protein create fewer problems.[17-20]

Once the 3D structures are defined, there is a need to assess the potential to generate a small molecule with the desired activity and selectivity. Such an assessment is based upon accurate knowledge of the key amino acids that modulate the binding site or

the active site. Computer-assisted molecular modeling (CAMM) links biological and structural data of the target and compound databases to generate ligands that possess structural and chemical complementarity to the target.[21-24] One key docking methodology —— docking of small molecules to protein binding sites were explored in the early eighties.[21] One of the first computer programs of this kind is DOCK, which was developed by Irwin Kuntz et al [25] and very quickly docking became an active research area. DOCK primarily explores three aspects of inhibitor design: creating negative image of the active site, placing the putative ligands into the sites, and evaluating the quality of the fit. Based on the treatment of ligand flexibility and protein flexibility, the docking methods can be divided into three categories: systematic methods, random or stochastic method and simulation methods. DOCK, as well as some other methods such as Flex-X,[26] Glide,[27,28] FLOG [29] belongs to systematic methods; AutoDock,[30] MOE-Dock [31] and GOLD [32] employ random method; and DOCK, Glide, AutoDock, MOE-Dock all include simulation method. Many other programs are also developed emphasizing on other objectives.[33,34] CAMM has become an indispensable tool in the drug discovery process and it leads to higher hit rate and often times higher ligand affinity.

Structure-based drug design is at its most power when it is coupled with combinatorial chemistry.[20,35] It enables the generation of a huge number of novel compounds by paralleling the synthesis of molecules. Its renaissance changed the way new chemical entities are generated. Combined with high-throughput screening (HTS), combinatorial chemistry dramatically accelerates the whole drug discovery process. The specific knowledge of the target makes the random combinatorial libraries evolve into

focused combinatorial libraries, and therefore access successful drug candidates more rapidly.

The lead compound identified from HTS was further modified and by going through the design cycle repeatedly, successful candidates are put into clinical trials for advanced pharmaceutical studies.

Structure-based drug design has already contributed to the discovery of many successful drugs in the market,[36] e.g. the peptidomimetic HIV-protease inhibitor Invirase™ (Hoffman-La Roche) and Norvir™ (Abbott); the neuraminidase-based antiinfluenza drug, Relenza™ (GlaxoSmithKline) and Tamiflu™ (Hoffman-La Roche); the anti-glaucoma carbonic anhydrase II (CA-II) inhibitor Trusopt® (Merck); Thrombin inhibitor Exanta™ (AstraZeneca). On the other hand, structure-based drug design is still imperfect. Uncertainties in the identity of location of the protein or ligand atoms, effects of crystallization conditions, and identification and location of water molecules derived from the electron density data can mislead medicinal chemist.[13] These ambiguities become quite significant when building an atomic model and they can therefore cause erroneous design strategies. At the same time, docking and scoring algorisms are not perfect.[22] The current calculation of ligand affinity is not precise enough to guide the design in the narrow range of affinity variations observed from the lead optimization. But facing a humongous pool of compounds and ever increasing crystal structures, the future for structure-based drug design has never been better. More excitingly, Professor David Baker's lab at University of Washington has developed a new algorism called "Rosetta". With the aid of this program, they have achieved great success in predicting protein structures and their interactions with high accuracy.[37-39] This landmark work could open

a new era in which computer modeling is making exceedingly important contributions to structural biology and drug discovery.

Substrate Analogue, Transition Analogue and Product Analogue

Structure-based drug design presents a lot of opportunities in drug discovery. Nevertheless, there are circumstances in which rational design of inhibitors can be performed without a target crystal structure. Take the development of protease (e.g. renin and HIV-1) inhibitor as an example, people proposed a two-step protocol using a substrate-based strategy for inhibitor design: i) characterize the substrate specificity of the enzyme and ii) synthesize peptides with similar features but replace the hydrolyzable amide bond with a nonreactive "isostere".[40,41] The peptide character of the inhibitor is subsequently reduced by modification of the side chains or the backbone. Inhibitors generated by this protocol are called "substrate analogue".

Another common design paradigm is transition state analogue. Transition state theory for enzyme-catalyzed reactions usually treats the transition state as a stable state and it proposes that tight-binding at the transition state is a critical feature of catalysis. Enzymatic transition states are dynamic and can not be captured by direct physical observations. Fortunately, the physical nature of the enzyme–substrate complex at the transition state can be approached experimentally in complexes of stable compounds that mimic the transition state.[42-45] Tightly bound transition state analogues can freeze the dynamic transition state of an enzyme then convert it into a thermodynamically stable state that resembles the transition state. The dynamic excursions of catalytic site loops, flaps and protein domain segments toward the transition state are portrayed and

immobilized at positions that mimic the transition state. The fidelity of the enzyme-transition state complex is determined by how well the transition state analogue mimics the actual reaction transition state. The most important part in the dynamic excursion is to establish the closest geometry mimic in the transition state. In most cases, substrate analogue and transition state analogues bind in a similar fashion, only transition state analogue causes enzyme to fold tighter around the catalytic site.[46,47]

Product analogue is also an analogue that is commonly studied. Substrate analogue and product analogues bound to the catalytic site to provide the proof of interactions at the reaction center. Product analogue is particularly useful in X-ray crystallography studies of the global and local conformational changes of the enzyme.[48]

Cholera Toxin, *Escherichia coli* Heat-labile Exterotoxin and AB_5 toxin family

Cholera is an acute diarrheal disease caused by the infection of small intestine with the bacterium *Vibrio chloerae*. Throughout history, populations all over the world have sporadically been affected by devastating outbreaks of cholera. There have been six recorded pandemics of cholera and we are in the midst of the seventh. Millions of cases have been reported and hundreds of thousands of deaths have been caused. In a recent outbreak in Angola, more than 1,200 people were killed in three months.[49,50]

Infected persons are usually characterized by moderate to profuse watery diarrhea, fall of blood pressure, vomiting, cramps in leg and abdomen and complete collapse.[51-53] In severe cases, rapid loss of body fluids leads to dehydration and shock. Without treatment, death can occur within hours of onset. People become infected by drinking water or eating food contaminated by the bacteria, mostly due to poor sanitation.[54] The

most important treatment for cholera is oral rehydration therapy (ORT), which consists of prompt replenishment of the water and salts lost in watery diarrhea and vomiting. ORT takes advantage of glucose-coupled sodium transport. The rehydration solution comprises of glucose and sodium and its composition is crucial for optimal absorption.[55] Although antibiotic such as Ciprofloxacin, Tetracycline and a recently discovered Azithromycin are used in cases of severe cholera, drug resistance problems for medicines against this diarrheal illness have been developing and the drugs are growingly less effective.[56] ORT thus remains the mainstay of cholera treatment. However, a major drawback of ORT is that it relies on the supply of large amount of clean water, which is usually extremely difficult in areas where outbreaks occur.

Intestinal fluid loss in cholera is primarily caused by the release of a heteromeric toxin —— cholera toxin (CT).[51-53] On release from the bacterium, CT binds to the epithelial enterocyte and it then gets internalized through the retrograde endocytosis.[57] Within the cell, CT constitutively activates adenylyl cyclase by activation of the stimulatory G protein G_{sa},[58] resulting in elevated levels of intracellular cAMP. The resulting increase in cAMP level activates ion channels, causing a greatly enhanced efflux of ions and water into the intestinal lumen and leading to a large volume of watery diarrhea. CT is composed of an A subunit and five B subunits: the catalytic activity resides in A and receptor binding function of the toxin is carried out by the B pentamer.[59]

Traveler and infants' diarrhea caused by enterotoxigenic *Escherichia coli* (ETEC) is reported to be responsible for 500,000 deaths worldwide in 1998 according to the World Health Organization (WHO).[60] Heat-labile enterotoxin (LT) is the major virulent factor for ETEC. LT is also an AB_5-type heterohexamer and it is closely related to CT.

They have about 80% sequence identity for both the A and B chains.[61] CT and LT belong to the cholera toxin family, which includes CT itself, *E. coli* heat-labile enterotoxin (LTs) LT-I and LT-II and a CT-like toxin from *Campylobacter jejuni*.[59, 61] Toxins of this family have a catalytically active A subunit toward ADP-ribosylation reaction and five identical B subunits for receptor binding. The same pentameric folding motif of the CT toxin family is shared by a diverse class of toxins, the so-called "bacterial AB_5 toxins".[59, 62-64] This AB_5 toxin family also includes Shiga toxin (SHT) family [65] and Pertussis toxin (PT),[66] which are from bacteria *Shigella dysenteriae* and *Bordetella pertussis* respectively. Shiga toxin family is made up of a number of toxins from *Shigella dysenteriae* and the shiga-like toxins [67] (SLTs, also known as Verotoxin) from certain *E. coli* strains (SLT-I, SLT-II), with SHT and SLT-I almost 100% identical. They have 60% sequence identity with SLT-II. Sharing very low sequence similarity, PT remarkably preserves the structural homology of CT and SHT families.[64]

Mechanism of Cholera Toxin Action

CT and LT are very similar both in structure and function. Both A and B subunits of CT are synthesized in the cytosol inside the bacterium cell.[68] A and B subunits of the toxin are then translocated across the inner membrane to the periplasm of the bacterium, where one A and five B subunits assembles into a hetero-hexameric AB_5 holotoxin. The toxins are secreted across the outer membrane by the extracellular protein secretion (Eps) and the general secretion pathway (Gsp) export apparatus of *V. cholerae* and ETEC, respectively, which are examples of the type II bacterial protein secretion system.[69] Once in the lumen of gastrointestinal tract of human host, B subunit recognizes and binds the

ganglioside GM1 receptor on the surface of human cell. This sets off the endocytosis and the toxin achieves its entry into the human cell. Through cell trafficking inside the cell, the toxin is first transported retrograde from plasma membrane to trans-Golgi.[57,70] The presence of endoplasmic reticulum (ER) targeting motif (C-terminal KDEL sequence) then indicates the delivery of the holotoxin to the ER.[57,71,72] Once in the ER, a portion of the A-subunit, the A1 chain, unfolds and separates from the B-subunit by interacting with protein disulfide isomerase (PDI).[73,74] The active A1 is instantaneously released by reduction and rapidly and spontaneously refolds. The A1 chain, which has been proteolytically cleaved off the A2 chain, is free to translocate through the SEC61 channel to gain access to the cytosol.[75,76] The B subunits proceed in the anterograde direction and travel to the lysosome for degradation.[70] Once in the cytosol, the A1 domain accesses its substrate, the alpha subunit of the stimulatory G protein ($G_{s\alpha}$), and modifies it through an NAD-dependent ADP-ribosylation reaction.[77] ADP-ribosylation of Arg201 of $G_{s\alpha}$ locks the G protein in its GTP-bound form and persistently stimulates the adenylate cyclase. The consequent dramatic production of cAMP activates the Cl^- channels.[70,78] The imbalance of ions causes an efflux of cellular water to the gut, which eventually produces host dehydration and diarrhea (**Figure 1.1**).

The Structure of Cholera Toxin

The crystal structure of CT holotoxin was first reported in 1995 by Zhang et al.[79] This AB_5 toxin is characterized as a catalytically active A subunit (27.2 kDa) and five identical B subunits (11.6 kDa) forming a symmetrical pore-like B pentamer.[62] The A subunit sits on top of the B pentamer. It has two distinctive functional components: a

wedge-shaped A1 domain, containing the enzyme-active site and an elongated A2 domain consisting of an α-helix with a "tail" that extends through the pore formed by the B pentamer.[80] The interaction between A and B subunits is mostly modulated by the A2 domain and the B pentamer. **Figure 1.2** is a representation of the structure of an enzymatically inactive CT holotoxin. Activation of the toxin requires a critical post-translational modification. This process involves severing the covalent connections between A1 and A2, which is achieved by proteolytic cleavage of the peptide loop and reducing the disulfide bond between these two domains.[81] The GM1-binding site is found on the "bottom" of each B subunit of the B pentamer opposing the A subunit.[82] The holotoxin structure reveals three potential target areas to design CT therapeutic agents. These include blocking the enzyme active site located in the A subunit, disrupting the assembly of the holotoxin by interrupting the A2–B interaction and intercepting the receptor binding to the bottom of the pentamer. Among these possible approaches, blocking the enzymatic reaction hasn't been actively pursued. As a consequence, the work described in chapter 2 will exclusively concentrate on this subject. Descriptions of specific strategies on the inhibitor design of receptor binding and toxin assembly have been described elsewhere.[83-87]

Activation of Enzymatic Activity

It has always been a challenge to unravel the molecular mechanism of the ADP-ribosylating activity of CT. The major hurdle is that only the structure of inactive form of wild-type CT holotoxin is available.[79] Thus very limited information about the active site cleft and substrate binding mode can be extracted. In spite of these difficulties, it has

been known that the proteolytic cleavage of the A subunit between residues192 and 195 is essential for the activation of the toxin.[81] Meanwhile, an alternative approach of studying various mutants of A subunit of LT has been established. By utilizing this indirect method, a partial model of activating the toxin's catalytic activity has been developed.[81] Amongst these efforts, the Arg7Lys mutant is especially enlightening. The structure of this mutant is fundamentally the same as wild-type LT, except that the A: 47-56 loop is mobilized. Even though the substitution of arginine to lysine is generally considered conservative, it results in markedly different interactions and therefore allows the active site crevice to be more exposed. These discoveries directly lead to the proposal by van den Akker et al of activating the A1 subunit through a series of conformational changes. This process is initiated by proteolytic nicking and disulfide bond reduction that occur at over 20 Å away from the catalytic site. This hypothesis pointed out that several residues could be critical for the catalytic function of CTA.[81,87] More importantly, these findings elicit extensive systematic mutagenesis research, which ultimately leads to the breakthrough of recruiting CTA1:Y30S variant as a replacement for the wide-type CT for further investigation of the enzymatic reaction.[80]

Reaction Mechanism of ADP-ribosylation

Both CT and LT covalently modify the alpha subunit of stimulatory G protein. The residual Arg201 of $G_{s\alpha}$ is of particular importance in this ADP-ribosylation reaction. It modifies the substrate nicotinamide adenine dinucleotide (NAD), prompting the release of the nicotinamide.[88] In the process of this transformation, the stereochemistry of the *C1'* has been inverted, suggesting an S_N2 mechanism (**Figure 1.3**). Mutational studies of

the active site identified three residues to be of great importance to the catalysis — Glu110, Glu112 and His44. Glu110 has been proved to be essential, mutation of this residue deteriorated CT's ADP-ribosyltransferase activity.[89,90] Glu112 has been proved by numerous biochemical studies as an absolute requirement for the catalytic reaction and it is also strictly conserved in all ADP-ribosylating toxins.[89,91] Both glutamates are required for full enzymatic activity.[92] His44 is the other critical residue that has been extensively studied. Mutating this histidine to an asparagine results in a complete loss of ADP-ribosylation activity of the $G_{s\alpha}$.[89,91] A His44Tyr mutation was reported to destabilize CT holotoxin.[92] It has been proposed that His44 is involved in the activation of the acceptor guanidine moiety by making a hydrogen bond to increase its nucleophilicity.[91] Glu112, the catalytically important residue has been proposed to stabilize the oxocarbonium. This transation state would then weaken the N-glycoside bond, making it susceptible to the nucleophilic attack by arginine. The His44-activated arginine would then cleave the N-glycosidic bond by an S_N2 mechanism, releasing nicotinamide and ADP-ribosyl-arginine.[81] However, recent kinetic isotope effects measured by Schramm's group predict that the ADP-ribosylation is mediated through a ribose oxocarbenium transition state via a S_N1-like mechanism.[93-95] Despite the discrepancy of the reaction mechanism, it will not alter our design of CT inhibitors to compete with the natural substrates in binding to CT because CT has binding sites for both substrates.

Scope of Thesis

The work described here in this thesis is divided into two parts: part 1 elucidates inhibitor studies for cholera toxin; part 2 is committed to the discovery of α-helix mimetics to probe the dimer interface of dormancy survival regulator of tuberculosis using a structure-based design methodology. Different strategies are implemented because of the restrictions on the structural information. In the cholera toxin project, crystal structure of active toxin is unavailable — this brings forth the employment of the substrate analogues. In the case of dormancy survival regulator (DosR) project, X-ray crystal structure clearly outlined the interactions and residues involved, enabling structure-based inhibitor design.

Chapter 1 is an outlook of structure-based drug design with some of the latest development in this field. Background on cholera, cholera toxin and its action is also enclosed. However, cholera toxin's structure and mechanism of its enzymatic reaction is what I intent to center on, because both are closely related to the cholera toxin inhibitor design.

Chapter 2 describes our inhibitor design effort on cholera toxin, which utilizes the integration of mechanism-based drug design and substrate analogue design to create potent inhibitors targeting CTA1. It demonstrates that rational design and efficient optimization of ligand affinity can be achieved by combining these two methodologies. The natural substrates NAD and arginine of $G_{s\alpha}$ provides an excellent starting point for a bisubstrate analogue with a higher affinity. The nicotinamide analogue and the guanidine moiety of arginine201 have been retained to avoid possible disruption of the existing

favorable interactions. The work described in Chapter 2 includes the design principles, chemical synthesis, spectroscopic characterizations, and kinetic studies of CTA inhibitors.

Chapter 3 is a brief review of the background of protein–protein interfaces and their corresponding interactions. Some typical systems are listed and some general strategies to address this class of interactions are summarized.

In chapter 4, our attempts in search of α-helix mimetics were executed on the basis of a 2.0 Å crystal structure of DosR C-terminal domain. We take two different approaches to access the problem: peptide approach and non-peptide small molecule approach. The small molecule approach yielded very potent inhibitors, yet the result was not very conclusive due to potential protein precipitation. Within the peptide approach, our objective is to find peptides that preferentially adopt an α-helix conformation. This goal is obtained by either optimizing the native sequence of the α10 helix or exerting external constrains to force the peptides into the preferred conformation. Mutations of the native sequence and side chain cyclization are among the methods we tried. Mass spectrometry, two-dimensional NMR, and circular dichroism are adopted to characterize the identity and α-helical content of the peptides. The synthesized peptides are evaluated using a gel-shift binding assay. Finally, some conclusions are drawn at the end of each project.

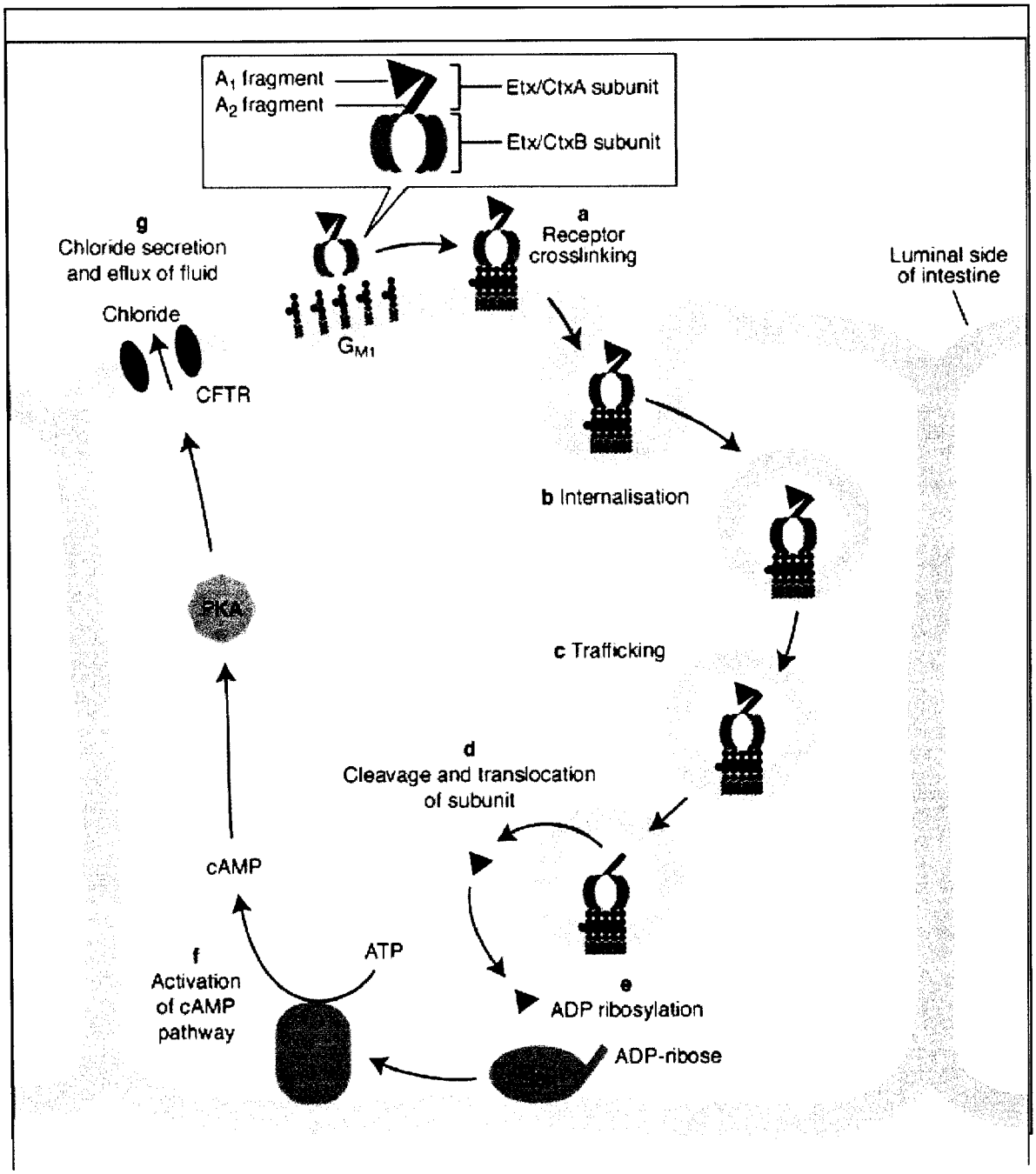


Figure 1.1. The invasion mechanism of cholera toxin in the infected host cell. This figure is adapted from ref. 78.

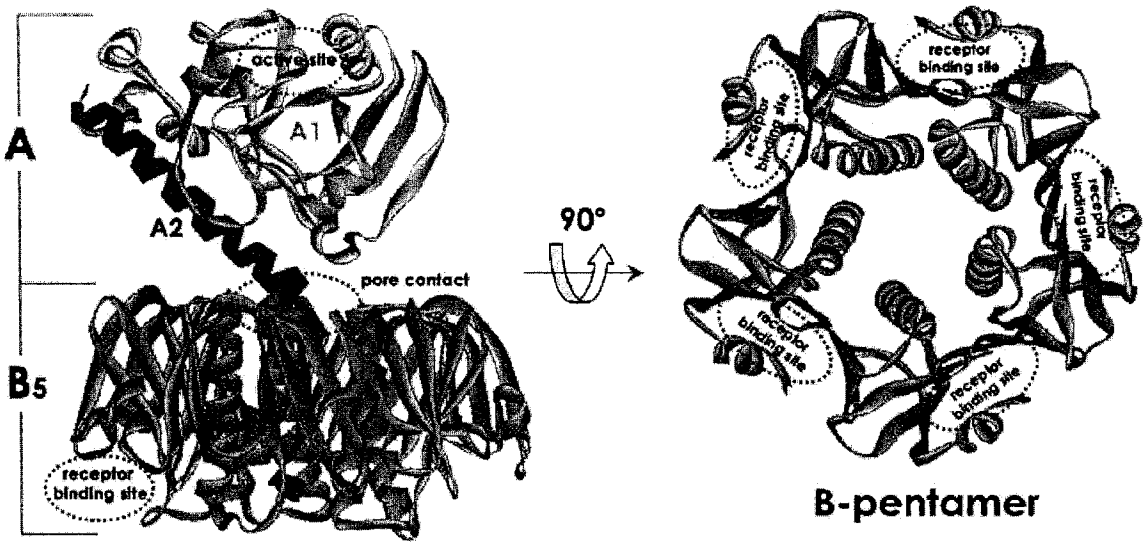


Figure 1.2. Crystal structure of cholera toxin holotoxin (left). The holotoxin is composed of one A subunit (top) and five B subunits (bottom). A subunit can be subdivided into A1 and A2, with A2 tethering A and B5 together. The active site, located in A1, carries the catalytic function. B5 is in charge of receptor binding. On the bottom of B pentamer, there are five copies of GM1 ganglioside binding sites (right). This figure was adapted from ref. 62 with the permission of the author.

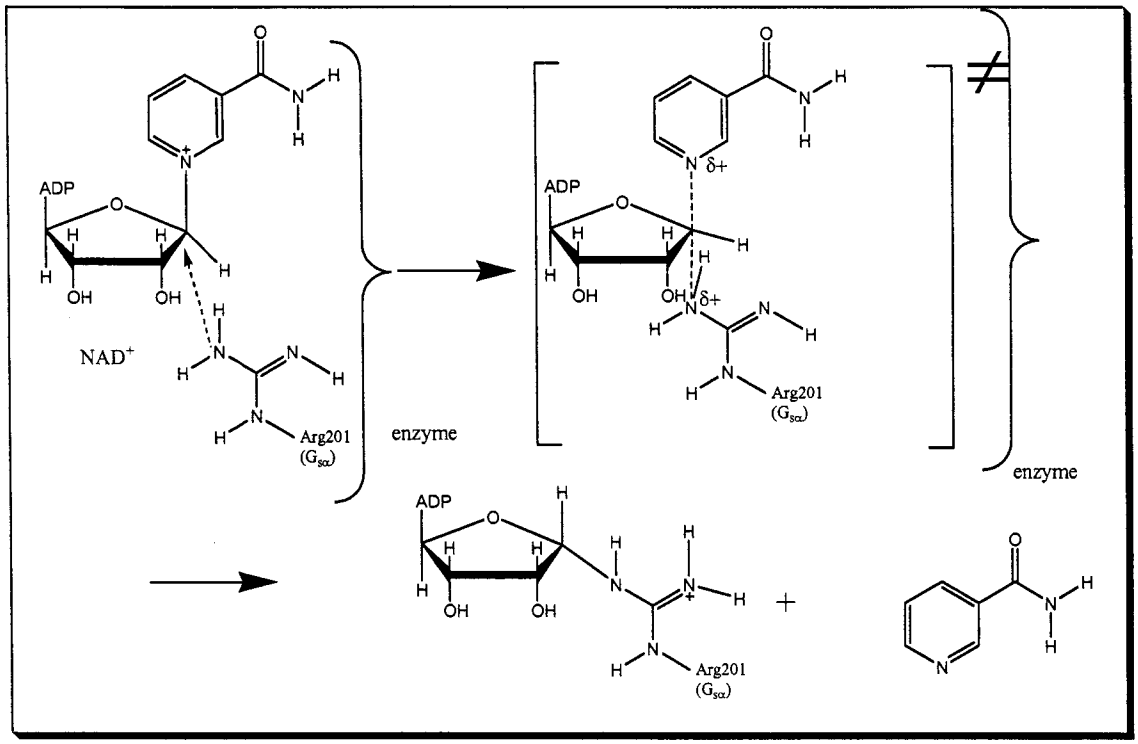


Figure 1.3. Reaction mechanism proposed for ADP-ribosylation that is catalyzed by CT. Kinetic study reveals that it may follow a S_N2 mechanism.

Notes to Chapter 1

1. Kuntz I. D. Structure-based strategies for drug design and discovery. *Science* **1992**, 257, 1078-1082.

2. Whittle P. J., Blundell T. L. Protein Structure-based drug design. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, 23, 349-375.

3. Hutchison C. A. III, Phillips S., Edgell M. H., Gillam S., Jahnke P., Smith M. Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.* **1978**, 253, 6551-6560.

4. Smith M., Leung D. W., Gillam S., Astell C. R., Montgomery D. L., Hall, B. D. Sequence of the gene for iso-1-cytochrome c in Saccharomyces cerevisiae. *Cell* **1979**, 16, 753-761.

5. Smith M. Site-directed mutagenesis. *Trends in Biochem. Sci.* **1982**, 7, 440-442.

6. Winter G., Fersht A. R., Wilkinson A. J., Zoller M., Smith M. Redesigning enzyme structure by site-directed mutagenesis: tyrosyl tRNA synthetase and ATP binding. *Nature* **1982**, 299, 756-758.

7. Smith M. In vitro mutagenesis. *Annu. Rev. Gen.* **1985**, 19, 423-462.

8. Liang N., Mauk A. G., Pielak G. J., Johnson J. A., Smith M., Hoffman B. M. Regulation of interprotein electron transfer by residue 82 of yeast cytochrome c. *Science* **1988**, 240, 311-313.

9. Smith M. The role of chemical synthesis in establishing gene function by in vitro mutagenesis. *Proceedings of the Robert A. Welch Foundation Conference on Chemical Research.* **1985**, 29, 439-455.

10. Clore G. M., Gronenborn A. M. Structures of larger proteins in solution: three- and four-dimensional heteronuclear NMR spectroscopy. *Science* **1991**, 252, 1390-1399.

11. Wuthrich K. Protein structure determination in solution by nuclear magnetic resonance spectroscopy. *Science* **1989**, 243, 45-50.

12. Hol W. G. Protein crystallography and computer graphics—toward rational drug design. *Angew. Chem. Int. Ed. Engl.* **1986**, 25, 767-778.

13. Davis A. M., Teague S. J., Kleywegt G. J. Application and limitations of X-ray crystallographic data in structure-based ligand and drug design. *Angew. Chem. Int. Ed. Engl.* **2003**, 42, 2718-2736.

14. Kuhn P., Wilson K., Patch M. G., Stevens R. C. The genesis of high-throughput structure-based drug discovery using protein crystallography. *Curr. Opin. Chem. Biol.* **2002**, 6, 704-710.

15. Kuhn P., Wilson K., Patch M. G., Stevens R. C. The genesis of high-throughput structure-based drug discovery using protein crystallography. *Curr. Opin. Chem. Biol.* **2002**, 6, 704-710.

16. Gwaltney S. L. II Design and synthesis of potent, selective, and orally efficacious DPP4 inhibitors accerated by high-throughput structural biology. *30th National Medicinal Chemistry Symposium*, June 25-29, Seattle, **2006**.

17. Wuthrich K. NMR of proteins and nucleic acids. Wiley, New York, **1986**.

18. Pellecchia M., Sem D. S., Wuthrich K. NMR in drug discovery. *Nat. Rev.* **2002**, 211-219.

19. Wishart D. NMR spectroscopy and protein structure determination: Applications to drug discovery and development. *Curr. Pharm. Biotech.* **2005**, 6, 105-120.

20. Henry C. M. Structure-based drug design. *Chem. Eng. News* **2001**, 79, 69-74.

21. Kuntz I. D., Blaney J. M., Oatley S. J., Langridge R., Ferrin T. E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **1982**, 161, 269-288.

22. Kitchen D. B., Decornez H., Furr J. R., Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* **2004**, 3, 935-949.

23. Shoichet B. K., Kuntz I. D., Bodian D. L. Molecular docking using shape descriptors. *J. Comp. Chem.* **1992**, 13, 380-397.

24. Meng E. C., Shoichet B. K., Kuntz I. D. Automated docking with grid-based energy evaluation. *J. Comp. Chem.* **1992**, 13, 505-524.

25. Ulrich R. Dock around the clock—current status of small molecule docking and scoring. *QSAR & Comb. Sci.* **2006**, 25, 605-615.

26. Rarey M., Kramer B., Lengauer T., Klebe G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, 261, 470-489.

27. McNally V. A., Gbaj A., Douglas K. T., Stratford I. J., Jaffar M., Freeman S., Bryce R. A. Identification of a novel class of inhibitor of human and Escherichia coli thymidine phosphorylase by *in silico* screening. *Bio. Med. Chem. Lett.* **2003**, 13, 3705-3709.

28. Halgren T. A., Murphy R. B., Friesner R. A., Beard H. S., Frye L. L., Pollard W. T., Banks J. L. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, 47, 1739–1749.

29. Kearsley S. K., Underwood D. J., Sheridan R. P., Miller M. D. Flexibases: a way to enhance the use of molecular docking methods. *J. Comput. Aided Mol. Des.* **1994**, 8, 565–582.

30. Olson A. J., Goodsell D. S. Automated docking in crystallography: analysis of the substrates of aconitase. *Proteins* **1993**, 17, 1-10.

31. Chemical Computing Group. MOE. Montreal, Quebec, Canada, **2003**.

32. GOLD Version 1.2, [http://www.ccdc.cam.ac.uk/products/life_sciences/gold/], **2003**.

33. McMartin C., Bohacek R. S. QXP: Powerful, rapid computer algrithms for structure-based drug design. *J. Comput. Aided Mol. Design* **1997**, 11, 333-344.

34. ThistleSoft, Inc. FLO96 ThistleSoft, Colebrook, CT, **1996**.

35. Borman S. Drugs by design. *Chem. Eng. News* **2005**, 83, 28-30.

36. Williams S. P., Kuyper L. F., Pearce K. H. Recent applications of protein crystallography and strcture-guided drug design. *Curr. Opin. Chem. Biol.* **2005**, 9, 371-380.

37. Baker D. Prediction and design of macromolecular structures and interactions. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2006**, 361, 459-463.

38. Schueler-Furman O., Wang C., Bradley, P., Misura K., Baker D. Progress in modeling of protein structures and interactions. *Science* **2005**, 310, 638-642.

39. Ashworth J., Havranek J. J., Duarte C. M., Sussman D., Monnat R. J. Jr., Stoddard B. L., Baker D. Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature* **2006**, 441, 656-659.

40. Greenlee W. J. Renin inhibitors. *Pharm. Res.* **1987**, 4, 364-374.

41. Debouck C. The HIV-1 protease as a therapeutic target for AIDS. *AIDS Res. Hum. Retroviruses.* **1992**, 8, 153-164.

42. Schramm V. L. Enzymatic transition states and transition state analogues. *Curr. Opin. Struct. Biol.* **2005**, 15, 604-613.

43. Smyth T. P. Substrate variants versus transition state analogues as noncovalent, reversible enzyme inhibitors. *Bioorg. Med. Chem.* **2004**, 12, 4081-4088.

44. Schramm V. L. Enzymatic transition states: thermodynamics, dynamics and analog design. *Arch. Biochem. Biophys.* **2005**, 433, 13-26.

45. Schramm V. L. Enzymatic Transition State Poise and Transition State Analogues. *Acc. Chem. Res.* **2003**, 36, 588-596.

46. Ringia E. A. T., Schramm V. L. Transition states and inhibitors of the purine nucleoside phosphorylase family. *Curr. Topics Med. Chem.* **2005**, 5, 1237-1258.

47. Wolfenden R., Radzicka A. Transition-state analogues. *Curr. Opin. Struct. Biol.* **1991**, 1, 780-787.

48. Schramm V. L., Shi W. Atomic motion in enzymatic reaction coordinates. *Curr. Opin. Struct. Boil.* **2001**, 11, 657-665.

49. Dziejman M., Balon E., Boyd D., Fraser C. M., Heidelberg J. F., Mekalanos J. J. Comparative genomic analysis of Vibrio cholerae: genes that correlate with cholera endemic and pandemic disease. *Proc. Natl. Acad. Sci. USA* **2002**, 99, 1556-1561.

50. World Health Organization. Global epidemics and impact of cholera. [http://www.who.int/topics/cholera/impact/en/], **2006**.

51. De S. N. Enterotoxicity of bacteria-free culture-filtrate of *Vibrio cholerae. Nature* **1959**, 183, 1533-1534.

52. Dutta N. K., Panse M. W., Kulkarni D. R. Role of cholera a toxin in experimental cholera. *J. Bacteriol.* **1959**, 78, 594-595.

53. Finkelstein R. A., Norris H. T., Dutta N. K. Pathogenesis experimental cholera in infant rabbits. *J. Infect. Dis.* **1964**, 114, 203-216.

54. World Health Organization. Cholera cases and deaths reported to WHO, 2002, [http://www.who.int/emc/diseases/cholera/choltbld2002.html]. *WHO Weekly* **2002**.

55. Pulungsih S. P., Punjabi N. H., Rafli K., Rifajati A., Kumala S., Simanjuntak C. H. Lesmana M., Subekti D., Sutoto F. O. Standard WHO-ORS versus reduced-osmolarity ORS in the management of cholera patients. *J. Health Popul. Nutr.* **2006**, 24, 107-112.

56. Sack D. A., Sack R. B., Chaignat C. L. Getting serious about cholera. *N. Engl. J. Med.* **2006**, 355, 649-651.

57. Lencer W. I. Retrograde transport of cholera toxin into the ER of the host cells. *Int. J. Med. Microbiol.* **1994**, 293, 491-494.

58. Robinshaw J. D., Russel D. W., Harris B. A., Smigel M. D., Gilman A. G. Deduced primary structure of the alpha subunit of the GTP-binding stimulatory protein of adenylate cyclase. *Proc. Natl. Acad. Sci. USA* **1986**, 83, 1251-1255.

59. Merrit E. A., Hol W. G. AB_5 toxins. *Curr. Opin. Struct. Biol.* **1995**, 5, 165-171.

60. Widdus R. Public-private partnerships for health: their main targets, their diversity, and their future directions. *Bulletin of the World Health Organization.* **2001**, 79, 713-720.

61. Sixma T. K., Pronk S. E., Kalk K. H., Wartna E. S., van Zanten B. A. M., Witholt B. , Hol, W. G. J. Crystal structure of a cholera toxin-related heat-labile enterotoxin from *E. coli. Nature* **1991**, 351, 371-377.

62. Fan E., O'Neal C. J., Mitchell D. D., Robien M. A., Zhang Z., Pickens J. C., Tan X. J., Korotkov K., Roach C., Krumm B., Verlinde C. L., Merritt E. A., Hol W. G. Structural biology and structure-based inhibitor design of cholera toxin and heat-labile enterotoxin. *Int. J. Med. Microbiol.* **2004**, 294, 217-223.

63. Feil I. K., Platas A. A., van den Akker F., Reddy R., Merritt E. A., Storm D. R., Hol W. G. Stepwise transplantation of an active site loop between heat-labile enterotoxins LT-II and LT-I and characterization of the obtained hybrid toxins. *Protein Eng.* **1998**, 11, 1103-1109.

64. Fan E., Merritt E. A., Verlinde C. L. M. J., Hol, W. G. J. AB5 toxins: structures and inhibitor design. *Curr. Opin. Struct. Biol.* **2000**, 10, 680-686.

65. Fraser M. E., Chernaia M. M., Kozlov Y. V., James M. N. Crystal structure of the holotoxin from Shigella dysenteriae at 2.5 A resolution. *Nat. Struct. Biol.* **1994**, 1, 59-64.

66. Yang J., Mou J., Shao Z. Structure and stability of pertussis toxin studied by in situ atomic force microscopy. *FEBS Lett.* **1994**, 338, 89-92.

67. Stein P. E., Boodhoo A., Tyrrell G. J., Brunton J. L., Read R. J. Crystal structure of the cell-binding B oligomer of verotoxin-1 from *E. coli. Nature* **1992**, 355, 748-750.

68. Pugsley A. P. The complete general secretory pathway in gram-negative bacteria. *Microbiol. Rev.* **1993**, 57, 50-108.

69. Robien M. A., Krumm B. E., Sandkvist M., Hol W. G. Crystal structure of the extracellular protein secretion NTPase EpsE of Vibrio cholerae. *J. Mol. Biol.* **2003**, 333, 657-674.

70. Kahn R. A., Fu H., Roy C. R. Cellular hijacking: a common strategy for microbial infection. *Trend in Biochem. Sci.* **2002**, 27, 308-314.

71. Majoul I. V., Bastiaens P. I., Soling H. D. Transport of an external Lys-Asp-Glu-Leu (KDEL) protein from the plasma membrane to the endoplasmic reticulum: studies with cholera toxin in Vero cells. *J. Cell Biol.* **1996**, 133, 777-789.

72. Lencer W. I., Constable C., Moe S., Jobling M. G., Webb H. M., Ruston S., Madara J. L., Hirst T. R., Holmes R. K. Targeting of cholera toxin and Escherichia coli heat labile toxin in polarized epithelia: role of COOH-terminal KDEL. *J. Cell Biol.* **1995**, 131, 951-962.

73. Tsai B., Rodighiero C., Lencer W. I., Rapoport T. A. Protein disulfide isomerase acts as a redox-dependent chaperone to unfold cholera toxin. *Cell* **2001**, 104, 937-948.

74. Bastiaens P. I., Majoul I. V., Verveer P. J., Soling H. D., Jovin T. M. Imaging the intracellular trafficking and state of the AB5 quaternary structure of cholera toxin. *EMBO J.* **1996**, 15, 4246-4253.

75. Schmitz A., Herrgen H., Winkeler A., Herzog V. Cholera toxin is exported from microsomes by the Sec61p complex. *J. Cell Biol.* **2000**, 148, 1203-1212.

76. Teter K., Holmes R. K., Schmitz A., Herrgen H., Winkeler A. and Herzog V. Inhibition of endoplasmic reticulum-associated degradation in CHO cells resistant to cholera toxin, Pseudomonas aeruginosa exotoxin A, and ricin. Cholera toxin is exported from microsomes by the Sec61p complex. *Infect. Immun.* **2002**, 70, 6172-6179.

77. Tsai B., Rapoport T. A. Unfolded cholera toxin is transferred to the ER membrane and released from protein disulfide isomerase upon oxidation by Ero1. *J. Cell Biol.* **2002**, 159, 207-216.

78. Salmond R. J., Luross J. A., Williams N. A. Immune modulation by the cholera-like enterotoxins. *Expert Rev. Mol. Med.* **2002**, 1 October, 1-16.

79. Zhang R. G., Scott D. L., Westbrook M. L., Nance S., Spangler B. D., Shipley G. G., Westbrook E. M., *J. Mol. Biol.* **1995**, 251, 563-573.

80. O'Neal C. J., Amaya E. I., Jobling M. G., Holmes R. K., Hol W. G. Crystal structures of an intrinsically active cholera toxin mutant yield insight into the toxin activation mechanism. *Biochemistry* **2004**, 43, 3772-3782.

81. van den Akker F., Merritt E. A., Pizza M., Domenighini M., Rappuoli R., Hol W. G. The Arg7Lys mutant of heat-labile enterotoxin exhibits great flexibility of active site loop 47-56 of the A subunit. *Biochemistry* **1995**, 34, 10996-11004.

82. Merritt E. A., Kuhn P., Sarfaty S., Erbe J. L., Holmes R. K., Hol W. G. The 1.25 Å resolution refinement of the cholera toxin B-pentamer: evidence of peptide backbone strain at the receptor-binding site. *J. Mol. Biol.* **1998**, 282, 1043-1059.

83. Pickens J. C., Mitchell D. D., Liu J., Tan X., Zhang Z., Verlinde C. L., Hol W. G., Fan E. Nonspanning bivalent ligands as improved surface receptor binding inhibitors of the cholera toxin B pentamer. *Chem. Biol.* **2004**, 11, 1205-1215.

84. Pickens J. C., Merritt E. A., Ahn M., Verlinde C. L., Hol W. G., Fan E. Anchor-based design of improved cholera toxin and E. coli heat-labile enterotoxin receptor binding antagonists that display multiple binding modes. *Chem Biol.* **2002**, 9, 215-224.

85. Merritt E. A., Zhang Z., Pickens J. C., Ahn M., Hol W. G., Fan E. Characterization and crystal structure of a high-affinity pentavalent receptor-binding inhibitor for cholera toxin and E. coli heat-labile enterotoxin. *J. Am. Chem. Soc.* **2002**, 124, 8818-8824.

86. Fan E., Zhang Z., Minke W. E., Hou Z., Verlinde C. L., Hol W. G. High-affinity pentavalent ligands of Escherichia coli heat-labile enterotoxin by modular structure-based design. *J. Am. Chem. Soc.* **2000**, 122, 2663-2664.

87. van den Akker F., Merrit E. A., Hol W. G. Structure and function of cholera toxin and related enterotoxins. *In handbook of experiemental pharmacology – bacterial protein toxins.* Vol. 145, 109-131. Springer-Verlag, Berlin, **2000**.

88. Locht C., Antoine R. A proposed mechanism of ADP-ribosylation catalyzed by the pertussis toxin S1 subunit. *Biochimie.* **1995**, 77, 333-340.

89. Lobet Y., Cluff C. W., Cieplak W. Jr. Effect of site-directed mutagenic alterations on ADP-ribosyltransferase activity of the A subunit of Escherichia coli heat-labile enterotoxin. *Infect. Immun.* **1991**, 59, 2870-2879.

90. Cieplak W. Jr., Mead D. J., Messer R. J. and Grant C. C. Site-directed mutagenic alteration of potential active-site residues of the A subunit of Escherichia coli heat-labile enterotoxin. Evidence for a catalytic role for glutamic acid 112. *J. Biol. Chem.* **1995**, 270, 30545-30550.

91. Kaslow H. R., Platter B., Takada T., Moss J., Mar V. L., Burnette W. N. Effects of site-directed mutagenesis on cholera toxin A1 subunit ADP-ribosyltransferase activity. *Bacterial protein toxins.* B. e. a. Witholt. Stuttgart, Jena, New York, Gustav Fisher. Zbl. Bakt. Suppl. **1992**, 23, 197-198.

92. Jobling M. G., R. K. Holmes. Biological and biochemical characterization of variant A subunits of cholera toxin constructed by site-directed mutagenesis. *J. Bacteriol.* **2001**, 183, 4024-4032.

93. Scheuring J., Schramm V. L. Pertussis toxin: transition state analysis for ADP-ribosylation of G-protein peptide alphai3C20. *Biochemistry* **1997**, 36, 8215-8223.

94. Scheuring J., Schramm V. L. Kinetic isotope effect characterization of the transition state for oxidized nicotinamide adenine dinucleotide hydrolysis by pertussis toxin. *Biochemistry* **1997**, 36, 4526-4534.

95. Rising K. A., Schramm V. L. Transition state analysis of NAD^+ hydrolysis by the cholera toxin catalytic subunit. *J. Am. Chem. Soc.* **1997**, 119, 27-37.

Chapter 2

Design and Structure-activity Relationship of Substrate Analogues as Cholera Toxin Inhibitors

Introduction and Summary

When we started this project, the only available crystal structure for cholera toxin (CT) holotoxin is in its inactive form. There were very few efforts towards the discovery of small molecular inhibitors that block the enzymatic function of the A subunit (i.e. the A1 chain). It is only until recently, Schramm's group from Albert Einstein College of Medicine, published their design of a series of transition state analogues on the basis of the stabilization of the oxocarbenium ion character of the transition states.[1] We sought to design some small molecules based on the mechanism of the ADP-ribosylation reaction (ADPR), which involves both the ribose donor and ribose acceptor, namely, nicotinamide adenosine dinucleotide (NAD) and Arg201 of α subunit of the stimulatory heterotrimeric G protein ($G_{s\alpha}$). It is our hypothesis that incorporation of the key moieties of both substrates would possibly enhance the affinity by providing stronger interactions with CT than its natural substrates. By binding CT's active site selectively, the compounds should be able to block its activity. We found that this simple modification of the natural substrate resulted in significant improvement in blocking CT in its enzymatic activity.

Results and Discussions

Inhibitor Design. Cholera toxin, an ADP-ribose transferase, carries out its catalytic function by transferring ADP-ribose of NAD to the arginine moiety of $G_{s\alpha}$,

resulting in the release of nicotinamide. The S_N2 mechanism suggested the involvement of both NAD and the arginine 201 of $G_{s\alpha}$.[2-4] On the basis of these data, we propose the design of a series of bisubstrate analogues that incorporate both the nicotinamide analogue and guanidine-containing moieties. To preserve the well-established interactions between substrate and enzyme, we replaced the nicotinamide with a benzamide; on the other end of the molecule, we simply changed arginine201 into an N-substituted guanidine derivative. An aliphatic linker is also included to allow systematic probe of the optimal spacing between the two moieties. By putting a variety of R groups with different properties on the N′ position of the guanidine (**Figure 2.1**), we enlarged the diversity and successfully built a library of over 20 compounds as the CTA inhibitors.

Library Synthesis. The guanidine library synthesis was performed on the solid support with a "TFA-Sensitive Arylsulfonylthiourea-Assisted Synthesis" methodology developed in our laboratory (**Scheme 2.1**). This synthesis can also be accomplished through a solution phase version of this method. The distinct advantage of our investigation is to achieve concomitant cleavage of product and removal of the activating arylsulfonyl group in solid-phase synthesis, without any loss in overall synthetic efficiency. Our method does not leave a guanidine protecting group after TFA cleavage, the side product was readily trapped and removed with simple aqueous extraction and filtration, which is advantageous when water soluble guanidinium compounds are desired.[5] Compared with other guanidinylation methods,[6,7] our method requires shorter reaction time and milder conditions. More importantly, our method produced higher isolated yields even for sterically hindered aliphatic amines or electron-deficient

arylamines. 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) also asserts its potential advantages over other TFA-sensitive arylsulfonyl activating groups such as Mtr. **Figure 2.2** summarizes the structures of the first generation inhibitors.

Assay Development and Inhibition Studies. The kinetic assay of CT was measured by using a high-performance liquid chromatography (HPLC)-based assay. This assay can monitor the progress of the CT catalyzed ADP-ribosylation reaction. The assay allows us to determine our compounds' inhibitory activity rapidly, quantitatively and reproducibly. The product, formed in the presence of NAD, can be well separated from the starting materials using HPLC and quantitated simultaneously (**Figure 2.3**). Interestingly, two peaks were present in an overlapping fashion close to the retention time of the product. In an attempt to further purify the product, we collected two peaks separately. The mass spectra and ^{1}H NMR spectra both matched very well, indicating them both being the product of the ADP-ribosylation reaction. In-depth analysis reveals that they were interchangable isomers generated in the course of the reaction.

To reduce the complexity of the assay system, we utilize an artificial substrate diethylamino(benzylidine-amino)guanidine (DEABAG) as an alternative for the natural ADP-ribose acceptor Arg201 of $G_{s\alpha}$ (**Figure 2.4**). DEABAG was first employed by Narayanan and colleagues in their effort to develop a fluorescence assay to quantitate the ADP-ribosylation product.[8,9] Compared with their assay, ours has superb reproducibility. An internal standard is essential in this assay for precise analysis. Theophylline, a water-soluble and UV-active compound, was selected as the internal standard. Its stability and good separation from the rest of the assay components ensures the consistency amongst

runs. The assay is the most sensitive when UV absorbance is detected at the optimum wavelength. Since the amount of the product is extremely small, it will be realistic to establish the optimum wavelength by measuring that of the starting materials. In fact, the use of DEABAG adds such distinct spectral characteristics to the product that it enables easy and quantitative detection. Wavelength scan for DEABAG, the product, and ADP (the chromophore of NAD) was conducted in the range of 200-400 nm. The optimum wavelength was determined to be 270 nm.

The exact amount of the product can be calculated by the integration ratio of the product peak and the internal standard peak, with calibration by a conversion factor from the extinction coefficient at 270 nm. When an inhibitor is added into the system, it binds selectively to CT. As a result, the effective concentration of CT drops and the product yield is lowered. By comparing the product formation with and without the inhibitors, we are able to derive the IC50 of our inhibitors over a wide range of concentration of the materials. The conversion factor of each constituent over the internal standard (theophylline) is derived from Beer's Law as the following.

Beer's Law: $A = \varepsilon \times b \times c$,

$A_{product} = \varepsilon_{product} \times b \times c_{product}$; $A_{theophylline} = \varepsilon_{theophylline} \times b \times c_{theophylline}$

$$\therefore \frac{A_{product}}{A_{theophylline}} = \frac{\varepsilon_{product} b c_{product}}{\varepsilon_{theophylline} b c_{theophylline}} = \frac{\varepsilon_{product} c_{product}}{\varepsilon_{theophylline} c_{theophylline}}$$

$$\therefore \frac{c_{product}}{c_{theophylline}} = \frac{A_{product} \varepsilon_{theophylline}}{A_{theophylline} \varepsilon_{product}} \quad \because A = \int_{t_2}^{t_1} peak \therefore \frac{c_{product}}{c_{theophylline}} = \frac{\int product \cdot \varepsilon_{theophylline}}{\int theophylline \cdot \varepsilon_{product}}$$

The conversion factors for NAD and DEABAG can be obtained in a similar manner:

$$\frac{c_{DEABAG}}{c_{theophylline}} = \frac{\int DEABAG \cdot \varepsilon_{theophylline}}{\int theophylline \cdot \varepsilon_{DEABAG}} \quad , \frac{c_{NAD}}{c_{theophylline}} = \frac{\int NAD \cdot \varepsilon_{theophylline}}{\int theophylline \cdot \varepsilon_{NAD}}$$

($\int NAD$ represents the integrated area of the NAD peak)

Once the extinction coefficients of NAD, DEABAG and Theophylline are experimentally measured, the conversion factor is resolved.

Library Screening and Kinetic Studies of Library Compounds. Each of the 19 first generation inhibitors in the guanidine library was screened. Glactose, a commercially available compound, was seleceted as a negative control (compound **6**). **Figure 2.5** shows that, compared with the control, these compounds display percentage inhibition ranging from ~10% to 98.8% (also see **Table 2.1**); among them, compounds **1**, **11**, **13**, **14** and **18**, in particular, exhibited apparent inhibition of over 90% at 1 mM in the enzymatic reaction. Full kinetic studies were carried out subsequently. In our investigation, ADP-ribosylation of DEABAG by CT was linear for 50 min. The initial rates were measured at various concentrations of DEABAG and NAD. The kinetics is very similar to those reported by Smets et al.[10] The Michaelis-Menten constants for both DEABAG and NAD were calculated from plots of the kinetic data.[11] The *K*m for NAD is 14.0 mM, which is very close to the earlier reported values.[6, 12-15] The apparent *K*m for DEABAG is 4.0 mM. Considering the fact that measuring K_i is by no means uncomplicated, we measured the IC50 of the library compounds instead. Percentage inhibition was plotted versus the logarithm of concentration, and the ensuing curve-fitting facilitates the IC50 determination. Under the assay conditions, those compounds showed

IC50 values of 125μM, 270μM, 200μM, 55μM, and 40μM respectively (see **Table 2.2**). Compound 18 stands out as the most potent inhibitor in the first generation.

Structure Activity Relationship and Inhibitor Optimization. From the marked differences in the inhibition, we gained some valuable insights about structure activity relationship (SAR). Although limited rather than thorough investigation has been done, we identified both benzamide (B) and the guanidino groups (G) are important for tight binding. Without any of the above two, significant loss in affinity is anticipated. No substitution on the benzamide ring was attempted assuming the benzamide moiety of the molecule is going to be positioned in the same hydrophobic pocket as the nicotinamide pocket in the natural substrate NAD. A tether was recruited to link the two portions together. Inhibitors with a single $-CH_2-$ linker between B and G demonstrated a consistent precedence towards the CT active site over those with no linker (**Table 2.1**). This result can be explained by the differences in their flexibility. For inhibitors without any linker, the molecules are very rigid and lack conformational flexibility; in addition, the direct linkage leads to the formation of a highly conjugated system, which literarily locks the compounds into a planar orientation. Meanwhile, when there is no linker in between, it ends up B and G being so close that it is no longer a good mimic of the substrates at transition state. On the contrary, inhibitors with a linker, even when it is a single methylene group, grant a lot more conformations (**Figure 2.6**). A linker of at least one carbon is required to position the hydrophobic moiety into the pocket.

Our SAR efforts were focused on the modifications of the R group. The removal of an isopropyl group resulted in a dramatic loss of potency (compound **3** vs **1**). The

replacement of a branched alkyl chain also resulted in significant loss of potency (compound **2** vs **1**). The binding pocket displays a great tolerance towards the aromatic substitutions on the guanidine. Replacing the phenyl ring with aliphatic rings resulted in a loss of potency. The removal of the bromine atom resulted in a compound with a comparable potency. In our second generation inhibitors, while we were trying to investigate the effect of different positioning of bromobenzyl, we also wanted to explore the CT binding pocket to a larger degree by introducing larger hydrophobic functionalities (**Table 2.1**). The *ortho* and *meta* substitutions of a bromine atom on the phenyl ring are tolerated. The introduction of a larger conjugated system did not improve the potency (for structures, see **Figure 2.7**). A linker between benzamide portion and guanidine portion of the inhibitor is definitely beneficial.. The introduction of hydrophilic substitutions (compound **9** vs **7**) resulted in a loss of potency. From the screening results and the IC50s of the initial compounds, we propose that the binding pocket is hydrophobic in nature. Compound **18**, which includes a 4-bromobenzyl moiety, is the most potent inhibitor. It demonstrated a 350-fold and 100-fold increase in potency compared with the natural substrate NAD and the artificial substrate DEABAG.

Cocrystallization of Library Compounds with CTY30S. From the mechanism of CT's reaction, it is obvious that the holotoxin molecule represents a pro-enzyme, and requires certain post-translational modifications, namely proteolytic nicking and disulfide-bond reduction, to generate the active A1 enzyme by separating it from the rest of the protein. The inactive holotoxin poses limitations for establishing relevance between existing models and inhibitor design approaches targeting CT, especially when

the structure of active A1 subunit is not available yet. Structures of other ADP-ribosylating toxins (ADPRTs) in complex with the substrate NAD are available and docking studies could be attempted based on these structures. However, it is possible that CTA1 utilizes an entirely different mechanism and/or substrate binding mode than these toxins. To fully apprehend the nature of the enzymatic reaction and to set the stage for future structure-guided inhibitor design, it is of great importance to obtain the crystal structure of substrate bound active CT.

In the past fifteen years, people have encountered a lot of challenges in obtaining activated CT. These obstacles were finally overcome by an Y30S mutant in the A subunit. The CTY30S toxin has the activity of activated wild-type enzyme without the activational modifications required by the wild-type holotoxin.[16] Because CTY30S is intrinsically active, it provides an excellent starting point for structural study of substrate- and inhibitor-binding to CT. Crystallographic experiments were conducted with small-molecule CT substrates and inhibitors using the active CTY30S toxin for cocrystallization, or pre-formed CTY30S crystals for soaking (by our collaborator in Professor Hol's laboratory). Although the poor affinity of CT for most substrates and inhibitors often prevented ordered binding in the CTY30S active site, certain CTY30S crystal forms that lacked crystal-packing constraints on the active-site loop could be soaked with high concentrations of NAD^+ to yield the first structures of substrate-bound CT. These structures confirm the identity of active site residues involved in substrate-binding and catalysis.

Most compounds, even when cocrystallized at 10 times the reported *K*m or IC50, were unable to displace the active-site loop. However, soaking high concentrations of

NAD^+ into crystal forms predisposed to active-site loop disorder did result in ordered substrate binding in two cases. Some key results and experimental conditions in the soaking study have been outlined in **Table 2.2**.

Crystal Structure of CTY30S (Glu110Asp, Glu112Asp) with NAD. Recently, Professor Hol's group has determined the structure of an inactive variant of CTA1 in complex with an activator protein, human ARF6. This system successfully overcame the obstacles that hindered structural study of substrate-binding using CTY30S, and structural details of a quaternary CTA1-NAD^+:ARF6-GTP complex were reported (**Figure 2.8**).[17] The binding mode disclosed in this structure builds a solid foundation for molecular modeling. The computer model was largely established with the coordinates that are determined in this structure.

Molecular Modeling. Docking studies were carried out by utilizing the software FLO96. This study was based on the crystal structure of a Glu112Asp, Glu110Asp double mutation of CTY30S (PDB: 2A5F).[17] To portray the active site interactions more accurately, we reversed the double mutation in our computer model and make them flexible and subject to conformational changes. In building this model, water molecules within the active site pocket were also considered as potential H-bond donor when defining the active site (i.e. HOH588, HOH643).

We docked benzamide into the binding site of NAD, the results showed most of the low energy conformers bound to the same pocket as NAD (**Figure 2.9**), which verified our initial hypothesis. This revealed that the presence of nicotinamide analogues

might contribute to the increase in affinity and selectivity. When we docked compound **18** into the cleft, we observed some minor changes in the binding mode as suggested by our model. The guanidino group formed hydrogen bonds with Glu110 and Glu112, pulling the molecule slightly out of the binding pocket (**Figure 2.10**). The additional hydrogen bonding, however, probably resulted in the affinity gain for compound **18**. As a comparison, docking results of compound **4**, in which guanidine and benzamide are directly connected, clearly showed the molecule either bound in a very different mode or it is directed more towards the solvent. This can be ascribed largely to the strain in the molecule. We also further investigated the potential possibilities of employing a longer tether between the benzamide fragment and guanidine moiety. Bearing this in mind, we docked compound **25** (**Figure 2.7**) with a two-carbon linker into the NAD binding pocket. To our surprise, the molecule totally abolished the binding mode of nicotinamide, indicating a completely different binding mode. Owing to the extra flexibility provided by the carbon chain, the guanidine portion of the molecule was bended to accommodate a hydrogen bond with Glu110 and Glu112 instead of Ala8 and Gln74, forming a likely π-stacking with the benzene ring. The benzamide portion was not in the nicotinamide binding pocket any more. The relatively small pocket in the vicinity of nicotinamide binding pocket has been fully occupied —— which could partially explain why we would not enhance the potency by simple substitution of the other side of the guanidines.

On the other hand, we have gained a lot of insights about future inhibitor design targeting CTA. According to the double mutant CTY30S structures, the diphosphate binding position has been confirmed. They formed quite a few H-bonds with binding pocket residues——these interactions are not sufficiently addressed in our current design

and are yet to be resolved hereafter. It is also noteworthy that the adenine ring makes a substantial contribution to the overall binding by forming efficient π-stacking with Arg25 and Arg11, which could provide feasible directions for future inhibitor design.

Assay Refinement. In our endeavor to differentiate inhibitors of different generations, we attempted to increase the sensitivity of the assay by decreasing the concentration of the inhibitors screened to 200 μM. There was no apparent affinity gain for any of the second generation inhibitors. In spite of the possible tolerance of substitutions on the guanidine, it occurred to us that it could also be caused by inhibitor promiscuity.[18] This elicited the characterization of our assay solution with dynamic light scattering (DLS). Analysis of some inhibitor solutions indicated aggregation at some level. This prompted us to refine our assay conditions.

Two methods were applied to address this problem. The first method we attempted was to introduce blocking additives such as bovine serum albumin (BSA, or other such additives include ovalbumine, casein etc.)[19-23] to the assay solution to eliminate non-specific binding. Take BSA as an example, we introduced low concentration of BSA into our assay system, and it stabilized our results. However, we lost some potency due to possible interactions between the inhibitors and BSA. Another solution we tried was to employ detergents to alleviate the aggregations, e.g. Triton X-100, Sodium dodecylsulfate, and so forth. The adverse effect was rather strong —— they either changed the kinetics dramatically or irreversibly deactivated the protein. Under current settings, it is very difficult to distinguish the difference between the second generation inhibitors and the lead compound conclusively. Although some research

groups have developed other methods to break the aggregation, such as adjusting pH, sonication, and pipetting,[24] etc. Unfortunately, they do not work in our case.

Conclusions

In this project, we aim to design small molecule inhibitors targeting the enzymatically active A subunit of cholera toxin. Without any crystal structure of the active enzyme, we designed a series of bisubstrate analogues based on the mechanism of the ADP-ribosylation reaction catalyzed by CT. Key features of the two involving substrate, namely NAD and Arg201 of Gs_{α}, are incorporated into the analogues. We developed a synthetic route to efficiently synthesize our focused guanidine library. Library compounds were screened against a CTY30S variant, which carries the full activity without extra activation. To evaluate our inhibitors, we developed an HPLC-based assay using DEABAG as the ADP-ribose acceptor. Good inhibiting compounds discovered in the screening were further investigated by kinetic studies. The most potent compound, **18**, demonstrated an IC50 of 40 μM. This is 350-fold more potent than the natural substrate NAD and 70-fold more potent than the artificial substrate DEABAG. We also performed an initial structure-activity relationship study, the results point out some properties of the binding site and some directions for future inhibitor design. With the recently-published crystal structure of a quaternary CTA1-NAD^+:ARF6-GTP complex, the interactions between the active site and NAD are unraveled for the first time. We built a computer model on the basis of the structure. In many cases, docking results successfully elucidates the interactions between inhibitors and the binding site and even provide reasonable explanations for the IC50s of some of the library compounds. Moreover, it sheds new light on designing CTA inhibitors with an improved affinity. A potential aggregation problem for some inhibitors is recognized and some measures have been taken to address this issue. Cocrystallization of CTY30S with some library

compounds was attempted by our collaborators in the group of Professor Wim G. J. Hol at the Biomolecular Structure Center of University of Washington.

Experimental Section

General. All chemical reagents were purchased from Aldrich, except Fmoc-Abz-OH was purchased from Advanced ChemTech, 4-methoxybenzylamine was purchased from Fluka and 2-phenylethylamine was purchased from Avocado Research Chemicals Ltd. Rink amide MBHA resin was purchased from Novabiochem. Poly-Prep chromatography columns for solid phase synthesis were purchased from Bio-Rad. All solvents used were obtained from Fisher Scientific. All reagents and common solvents were of reagent grade or HPLC grade and were used for synthesis and separation without further purification unless otherwise noted. Parallel syntheses were performed on a Qwest 210 synthesizer from Agonaut Technologies. Silica gel used in column chromatography was 70-230 mesh. HPLC purification was performed using a preparative Hewlett-Packard 1100 series HPLC system with a variable wavelength detector. The Columns used is a ZORBEX 300 SB-C18 reverse-phase column from Agilent (25 mm × 150 mm, 5 μm). ^{1}H-NMR spectra were recorded at 300 MHz on a Bruker AV300. Mass spectra were obtained using an Agilent 1100 series LCD/MS mass spectrometer equipped with an ion trap.

Focused Library and Precursor Synthesis.

Synthesis of Precursor 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf). A solution of 68 g (0.5 mol) of 2,3,5-trimethylphenol, 36 g (0.5 mol) of isobutylaldehyde, and 120 mL of toluene, and 6.5 g of concentrated sulfuric acid was allowed to reflux for overnight, with water being removed azetropically as it was formed through the use of a Dean-Stark distillation receiver. The reaction mass was then distilled at 20°C, collecting

the total distillate in a single receiver. The distillate was washed with 20% sodium hydroxide solution to remove unreacted phenol, and then with water, and dried over magnesium sulfate. It was then distilled again to give pure 2,2,4,6,7-pentamethyl courmaran product. 2,2,4,6,7-pentamethyl courmaran was then dissolved in anhydrous chloroform (500 mL) and cooled to -5°C. A solution of chlorosulphonic acid (140 mL, 2.10 mol) in dry chloroform (150 mL) was added maintaining the temperature of the reaction solution at ~0°C. After the addition was complete, the reaction was left to stir for 15 minutes at 0°C and another hour with cooling bath removed. The dark brown solution was then poured onto crushed ice and the organic layer separated and washed with stirred with 5% Na_2CO_3, saturated $NaHCO_3$, water and brine before drying over $MgSO_4$. After filtration, the bulk solvent was evaporated by rotary evaporator. The residual coumaran was then dissolved in either acetone or THF. A solution of $NH_3 \cdot H_2O$ (400 mL, 2.10 mol) in 400 mL acetone/THF was subsequently added to the diluted chlorosulfonylated coumaran keeping the temperature at 0°C. After finishing the addition of ammonium hydroxide, the reaction mixture was further stirring for 15 minutes at 0°C and an hour at room temperature. After filtration through a Buchner funnel, the filtrated was washed with water. It was then dried to give a product of 38.63 g. The crude product was then loaded onto a silicon gel column for flash chromatography. The column was eluted with 2.5% methanol in methylene chloride. The fraction of the final product was collected and evaporated (**Scheme 2.2**).

2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf). MS: 270.1 $(M+H)^+$. ^1H NMR (CD_3OD) δ1.42 (6H, s), 2.06 (3H, s), 2.44 (3H, s), 2.50 (3H, s), 2.98 (2H, s).

Solid-phase Synthesis of N,N'-substituted Guanidines. Rink amide MBHA resin (~0.1 g, loading 0.78 mmole/g) was soaked in N,N-dimethyl acetamide (DMA) for 30 minutes and then was deprotected with 20% piperidine in DMA for 30 minutes. After washing the resin thoroughly with DMA, solutions of 3-(N-Fmoc-aminomethyl) benzoic acid (5 eq.) or 3-(N-Fmoc-amino)benzoic acid (**1′**, **Scheme 2.1**, 5 eq.), PyBOP (5 eq.) and N,N-diisopropylethylamine (DIPEA, 10 eq.) in DMA were added to the resin. The reaction mixture was shaken gently overnight to produce resin-supported **A**. After washing, the resin loading was confirmed by the release of 9-methylene-9H-fluorene due to the Fmoc removal by 30 minutes treatment of 20% piperidine in DMA.

Reaction sequence 1: After Fmoc removal and washing, resin-supported **A** (~0.1 g) (or resin prepared from **1′**) was suspended in DMA with 10 eq. of DIPEA. To this was added compound **2′** (2.5 eq.) and the Mukaiyama reagent (2.5 eq.). The mixture was shaken gently overnight (12-18 hours). After washing, the resin was cleaved with a mixture of TFA:H_2O:TIS (94:3:3) for 1 hour. The crude product was extracted with water, filtered and purified by HPLC to give the final product.

Reaction sequence 2: After Fmoc removal and washing, resin-supported **A** (~0.1 g) (or resin prepared from **1′**) was suspended in 5 ml of methylene chloride. To this was added a solution of pentafluorophenyl thiochloroformate (5 eq.) and DIPEA (10 eq.) in methylene chloride. The reaction was allowed to proceed for 1 hour at room temperature with gentle shaking. The resin was then thoroughly washed with methylene chloride

followed by methyl sulfoxide (DMSO). Then a solution of $PbfNH_2$ (5 eq. to **A**) and potassium *tert*-butoxide (1 eq. to $PbfNH_2$ from 1 M THF solution) in DMSO was added to the resin. The reaction was shaken for 1 hour at room temperature, and then washed thoroughly with DMSO followed by N,N-dimethylformamide (DMF). Subsequently, the resin was treated with a solution of an amine nucleophile (10 eq. to **A**), DIPEA (7 eq.) and 1-[3-(Dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride (EDC, 5 eq.) in DMF for overnight at room temperature. After washing, the product was cleaved and purified as in "Reaction sequence 1".

Compound **2'** is synthesized using a solution phase method. This part of the work was carried out by previous members of the Fan lab and can be found elsewhere.[5]

Characterization of the N,N'-disubstituted Guanidine Library.

N,N-diisopropyl-N'-(3-aminocarbonylbenzyl)guanidine (1). From 0.081 mmole of resin-supported **A** and diisopropylamine, *1* was obtained in 76% yield as the TFA salt (23.0 mg). MS: 277.1 $(M+H)^+$. ^1H NMR (D_2O) δ1.15 (12H, d), 3.82 (2H, m), 4.41 (2H, s), 7.36 (2H, m), 7.56 (2H, m).

N,N-diethyl-N'-(3-aminocarbonylbenzyl)guanidine (2). From 0.094 mmole of resin supported **A** and diethylamine, *2* was obtained in 84% yield as the TFA salt (27.0 mg). MS: 249.1 $(M+H)^+$. ^1H NMR (D_2O) δ1.00 (6H, t), 3.23 (4H, qt), 4.36 (2H, s), 7.35 (2H, m), 7.55 (2H, m).

N-(tert-butyl)-N'-(3-aminocarbonylbenzyl)guanidine (3). From 0.081 mmole of resin-supported **A** and *t*-butylamine, *3* was obtained in 77% yield as the TFA salt (23.2

mg). MS: 249.1 $(M+H)^+$. ^1H NMR (D_2O) δ1.19 (9H, s), 4.35 (2H, s), 7.36 (2H, m), 7.60 (2H, m).

*N-(4-bromobenzyl)-N'-(3-aminocarbonylphenyl)guanidine (**4**).* From 0.078 mmole of resin prepared from **2'**, ***4*** was obtained in 78% yield as the TFA salt (28.0 mg). MS: 346.8 and 348.8 $(M+H)^+$. ^1H NMR (D_2O) δ4.31 (2H, s), 7.09 (2H, d), 7.30 (1H, d), 7.37-7.43 (3H, m), 7.48 (1H, s), 7.57 (1H, d).

*N-phenyl-N'-(3-aminocarbonylbenzyl)guanidine (**5**).* From 0.089 mmole of resin-supported **A** and aniline, ***5*** was obtained in 95% yield as the TFA salt (32.7 mg). MS: 269.1 $(M+H)^+$. ^1H NMR (D_2O) δ4.36 (2H, s), 7.42 (7H, m), 7.90 (2H, m).

N-(1-benzylpyrrolidin-3-yl)-N'-(3-aminocarbonylbenzyl)guanidine (7). MS: 367.2 $(M+H)^+$. ^1H NMR (CD_3OD) δ2.31 (2H, m), 2.77 (2H, m), 3.33 (3H, m), 4.40 (2H, s), 4.57 (2H, s), 7.59 (7H, m), 7.95 (2H, m).

*N-(2-(N-ethyl-N-(3-methylbenzenamine))aminoethyl)-N'-(3-aminocarbonylbenzyl) guanidine (**8**).* MS: 354.2 $(M+H)^+$. ^1H NMR (CD_3OD) δ1.21 (3H, t), 2.42 (3H, s), 3.52 (4H, m), 3.68 (2H, t), 4.48 (2H, s), 6.70 (1H, s), 6.96 (2H, d), 7.30(1H, t), 7.55 (2H, m), 7.94 (2H, m).

*N-4-(piperidinyl-1-ethyl-carboxylate)-N'-(3-aminocarbonylbenzyl)guanidine (**9**).* MS: 348.2 $(M+H)^+$. ^1H NMR (CD_3OD) δ1.30 (3H, t), 1.53 (2H, m), 1.83 (2H, m), 3.03 (4H, m), 3.69 (1H, m), 4.17 (2H, q), 4.55 (2H, s), 7.52(2H, m), 7.88 (2H, m).

*N-(2-(1-piperidinyl)-ethyl)-N'-(3-aminocarbonylbenzyl)guanidine (**10**).* MS: 304.2 $(M+H)^+$. ^1H NMR (CD_3OD) δ1.66-2.21 (6H, m), 3.09 (2H, m), 3.39 (3H, m), 3.62 (1H, m), 3.77 (2H, t), 4.62 (2H, s), 7.63 (2H, m), 7.93 (2H, m).

*N-(2-pehnyl)ethyl-N'-(3-aminocarbonylbenzyl)guanidine (**11**).* MS: 297.2 $(M+H)^+$. ^1H NMR (CD_3OD) δ3.53 (2H, t) 3.67 (2H, t) 4.54 (2H, s), 7.25-7.57 (5H, m), 7.62-7.74 (2H, m), 7.74-7.91 (2H, m).

*N-(4-methoxy)benzyl-N'-(3-aminocarbonylbenzyl)guanidine (**12**).* MS: 313.2 $(M+H)^+$. ^1H NMR (CD_3OD) δ3.46 (3H, s), 4.57 (4H, m), 7.11-7.35 (4H, m), 7.37-7.63 (2H, m), 7.75-7.83 (2H, m).

*N-(1-benzylpiperidine-3-yl)-N'-(3-aminocarbonylbenzyl)guandine (**13**).* MS: 366.2 $(M+H)^+$. ^1H NMR (CD_3OD) δ1.82 (2H, m), 2.17 (2H, m), 3.02 (2H, t), 3.51 (2H, m), 3.73 (1H, m), 4.55 (2H, s), 4.65 (2H, s), 7.42 (7H, m), 7.78 (2H, m).

*N-benzyl-N'-(3-aminocarbonylbenzyl)guandine (**14**).* MS: 283.2 $(M+H)^+$. ^1H NMR (CD_3OD) δ4.65 (4H, s), 7.07-7.21 (5H, m), 7.32 (2H, m), 7.76 (2H, m).

*N,N-diisopropyl-N'-(3-aminocarbonylphenyl)guanidine (**15**).* MS: 263.2 $(M+H)^+$. ^1H NMR (D_2O) δ1.66 (12H, d), 4.51 (2H, m), 7.42 (2H, m), 7.76 (2H, m).

*N-(tert-butyl)-N'-(3-aminocarbonylphenyl)guanidine (**16**).* From 0.105 mmole of resin prepared from **2'** and *t*-butylamine, *16* was obtained in 88% yield as the TFA salt

(32.0mg). MS: 235.0 $(M+H)^+$. ^1H NMR (D_2O) δ1.22 (9H, s), 7.27 (1H, d), 7.35 (1H, t), 7.45 (1H, s), 7.52 (1H, d).

N-phenyl-N'-(3-aminocarbonylphenyl)guanidine (17). From 0.105 mmole of resin prepared from **2'** and aniline, *17* was obtained in 70% yield as the TFA salt (27.2 mg). MS: 255.0 $(M+H)^+$. ^1H NMR (D_2O) δ7.16 (3H, m), 7.27 (2H, m), 7.36 (2H, m), 7.53 (2H, m).

N-(4-bromobenzyl)-N'-(3-aminocarbonylbenzyl)guanidine (18). From 0.078 mmole of resin-supported **A**, *18* was obtained in 82% yield as the TFA salt (30 mg). MS: 361.0 and 363.0 $(M+H)^+$. ^1H NMR (CD_3OD) δ4.46 (2H, s, CH_2), 4.55 (2H, s, CH_2), 7.31 (2H, d), 7.62 (4H, m), 7.95 (2H, m).

N-benzyl-N'-(3-aminocarbonylphenzyl)guanidine (19). MS: 269.1 $(M+H)^+$. ^1H NMR (CD_3OD) δ4.65 (2H, s), 7.22-7.40 (3H, m), 7.42-7.64 (4H, m), 7.82 (2H, m).

N-(1-benzylpiperidine-3-yl)-N'-(3-aminocarbonylphenzyl)guanidine (20). MS: 352.2 $(M+H)^+$. ^1H NMR (CD_3OD) δ1.82 (2H, m), 2.17 (2H, m), 3.02 (2H, t), 3.51 (2H, m), 3.73 (1H, m), 4.55 (2H, s), 7.40 (7H, m), 7.74 (2H, m).

N-(2-bromobenzyl)-N'-(3-aminocarbonylbenzyl)guanidine (21). MS: 361.0 and 363.0 $(M+H)^+$. ^1H NMR (CD_3OD) δ4.57 (4H, m), 7.20-7.51 (3H, m), 7.53-7.66 (3H, m), 7.70-7.90 (2H, m).

N-(3-bromobenzyl)-N'-(3-aminocarbonylbenzyl)guanidine (22). MS: 361.0 and

363.0 $(M+H)^+$. ^1H NMR (CD_3OD) δ4.41 (2H, s), 4.53 (2H, s), 7.20-7.43 (3H, m), 7.43-7.66 (3H, m), 7.80-7.90 (2H, m).

*N-(4-phenylbenzyl)-N'-(3-aminocarbonylbenzyl)guanidine (**23**).* MS: 359.2 $(M+H)^+$. ^1H NMR (CD_3OD) δ4.41-4.60 (4H, m), 7.23-7.52 (7H, m), 7.52-7.74 (4H, m), 7.74-7.91 (2H, m).

*N-(1-naphthyl)methyl-N'-(3-aminocarbonylbenzyl)guanidine (**24**).* MS: 333.2 $(M+H)^+$. ^1H NMR (CD_3OD) δ4.55 (4H, s), 7.38-7.74 (6H, m), 7.74-8.09 (5H, m).

N-(ADP-ribose)-N'-diethylamino(benzylidine-amino)guanidine (ADP-ribosylation reaction product in the assay). MS: 775.2 $(M+H)^+$. ^1H NMR (CD_3OD)δ1.04 (6H, t), 1.81 (4H, q), 3.93-4.55 (8H, m), 5.53 (1H, m), 6.05 (1H, d), 7.52 (2H, m), 7.85 (2H, m), 7.86-8.17 (1H, m), 8.18-8.27 (1H, m), 8.30-8.51 (1H, m).

Inhibition Studies and Compound Screening

General Assay Procedure. The assays are based on ADP-ribosylation of NAD^+ catalyzed by CT (**Figure 1.3**). All assays were initiated by the addition of CTY30S solution. Solutions of buffer PBS, NAD^+, Theophylline, and DEABAG were added in the exact order as listed above. The enzymatic assay was performed at 37°C and pH 7. Aliquots (25 μL) were taken from the reaction mixtures at designated times and quenched with 75 μL of 1% TFA. 20 μL of the quenched sample was injected into a Hewlett-

Packard 1100 series analytical HPLC system with a variable wavelength detector for analysis immediately. The product was quantitied by integration of its UV peak as compared to the peak of internal standard Theophylline. Inhibitors were competitive inhibitors with respect to NAD^+, and all kinetic data were fit to equations for competitive inhibition using the GraphPad Prism program (version 4). Values of *K*m and k_{cat} were determined by fits to the equation $v = k_{cat} ES/(S + K\text{m})$, where v is the initial rate, E is the enzyme concentration, S is the substrate concentration, ES is the enzyme-substrate complex concentration, k_{cat} is the catalytic turnover number, and *K*m is the Michaelis-Menton constant. IC50, the half maximal inhibitory concentration, is calculated by GraphPad Prism as well. Inhibition assays of CTA for its ADP-ribosylation were conducted using its mutant CTY30S. Solutions of all assay constituents were neutralized to pH 7 using standard HCl or NaOH solution before mixing.

Assay for CTAY30S. Active enzyme CTY30S mutant was kindly provided by Claire J. O'Neal as 4.5 mg/mL solution in buffer G at pH 7.4. The enzyme was reconstituted to 1.0 mg/mL (~37 μM) to stabilize. The activity was assayed at 37°C in 200 mM potassium buffer saline (PBS), at pH 7.0. The CTY30S was incubated at 37°C for 30 min before mixing with other ingredients of the assay.

Determination of *K*m and k_{cat} of ADP-ribosylation of DEABAG by CTY30S. Reaction mixtures of 500 μL contained 10.0 μM of CTY30S, 40 mM of NAD^+, 0.08 mM of Theophylline and variable concentrations of DEABAG (0.25, 0.5, 1.0, 1.5, 2.0, 4.0 mM of DEABAG). Aliquots (25 μL) were taken from the reaction mixture at 20, 50, 75,

100, 135 min and applied to HPLC for analysis for lower concentration; aliquots for higher concentration of DEABAG were taken from the reaction mixture at 5, 10, 15, 20, 30, 40, 50 min and applied to HPLC for analysis. To enhance the solubility, 1% of DMSO was used to help dissolve DEABAG in the stock solution. Since the stock solution was further diluted in the assay solution, the effect of DMSO should be ignored.

Determination of *K*m and k_{cat} of ADP-ribosylation of NAD^+ by CTY30S. Reaction mixtures of 500 μL contained 10.0 μM of CTY30S, 10 mM of DEABAG, 0.08 mM of theophylline and variable concentrations of NAD^+ (1.0, 3.0, 10.0, 20.0, 25.0 mM of NAD^+). Aliquots (25 μL) were taken from the reaction mixture at 10, 20, 30, 45, 60, 75, 100, 120 min and applied to HPLC for analysis.

Screening of CTA Inhibitors for ADP-ribosylation Reaction by DEABAG. To the CTY30S (10.0 μM) were added variable concentrations of the inhibitors (0.012, 0.037, 0.11, 0.33, 1.0 mM), 10.0 mM of NAD^+, 0.08 mM of theophylline, and 2 mM of DEABAG in a final volume of 500 μL. Aliquots were taken from the reaction mixtures at 120 min and applied to HPLC for analysis. Initial rates for these reactions were maintained for at least 3 hours.

IC50 Measurements. The CTY30S was performed as kindly provided by Claire O'Neal as previously reported.[16] Test samples consisted of 23.4 nM CTY30S preincubated with NAD^+ for 30 min at room temperature. Independent experiments for each inhibitor were carried out in triplicates and validated against an inhibitor

concentration gradient of 0.012, 0.037, 0.11, 0.33, 1.0 mM. IC50 values were calculated from triple data sets of at least five different concentrations of competitive inhibitors by nonlinear regression,[25] and reported as a weighted average of three independent determinations.

Dynamic Light Scattering. Protein or inhibitor in phosphate buffered saline (150 mM NaCl/15 mM phosphate, pH 7.4) was individually filtered through inorganic membrane filters (Whatman, Anodisc13, 0.02 μm) into separate vials. Then various portions of protein solutions themselves and their mixtures with inhibitors were, without further filtration, transferred into a sample cell for measurement. Dynamic light scattering (DLS) measurement was carried out on a DynaPro99 instrument (Protein Solutions Inc.) illuminated by a 25 mW, 832.8-nm-wavelength, solid-state laser at 20°C. Data analysis was performed using the dynamics Version 5.25.44 software provided with the instrument. Inverse Laplace transform ("regularization fit") analysis was used to find the mean and standard deviation (polydispersity) of the hydrodynamic radius (*R*h) distribution for the Inhibitor/protein complex species in solution.

Blocking Additives. Additives introduced into the assay are divided into two catagories: the first category is non-specific binding blockers, for example, BSA, ovalbumin, casein, chicken egg albumin. A series of scanning was done to decide which one causes the least amount of disturbances to the kinetic profile. BSA was chosen and subsequent optimization showed 0.2% of BSA gave the best results. The assay setup was almost the same as previously stated, except 100 μL of 0.2% BSA was introduced and the volume of PBS was changed accordingly. Solutions were added strictly in the order as

described above and the mixture was incubated at room temperature for 30 min after the addition of BSA. Aliquots were taken from the reaction mixture at 45, 90, 135 min and applied to HPLC for analysis.

Another category is surfactants, including sodium dodecyl sulfate, octyl glucopyranoside, CHAPS, glycoside, and Triton X-100. Triton X-100 gave a kinetic profile that is closest to that of the original assay, thus it was selected for further study. After serial scanning, 0.01% Triton X-100 was used. 100 μL was added after Theophylline. Aliquots were taken from the reaction mixture at 45, 90, 135 min and applied to HPLC for analysis.

Protein Production and Crystallographic Studies. CTY30S was expressed in *E. coli,* using constructs kindly provided by the Holmes lab from University of Colorado, successful expression, purification, and crystallization is conducted by the Hol lab as described in O'Neal et al.[16,17] Crystallographic studies about the quaternary CTA1-NAD^+:ARF6-GTP complex and cocrystallization of CTY30S with lead compounds were also carried out in the Hol lab.

Computer Modeling with FLO96. Docking was carried out by using the MCDOCK conformational searching by energy minimization procedure of FLO96.[26,27] Each compound was subjected to at least 1,000 cycles of conformer search and energy minimizations to identify conformations most likely to bind to the enzymatic pocket defined. Each cycle involved 400 rapid Monte Carlo search steps followed by energy minimization of the best conformer from the set of 400. The program returned the 20 lowest energy conformations found for visual inspection. The following criteria were

applied when evaluating the docked molecules: steric fit, hydrophobic interactions, hydrogen bonding, low molecular mechanics energy, and low internal inhibitor energy.

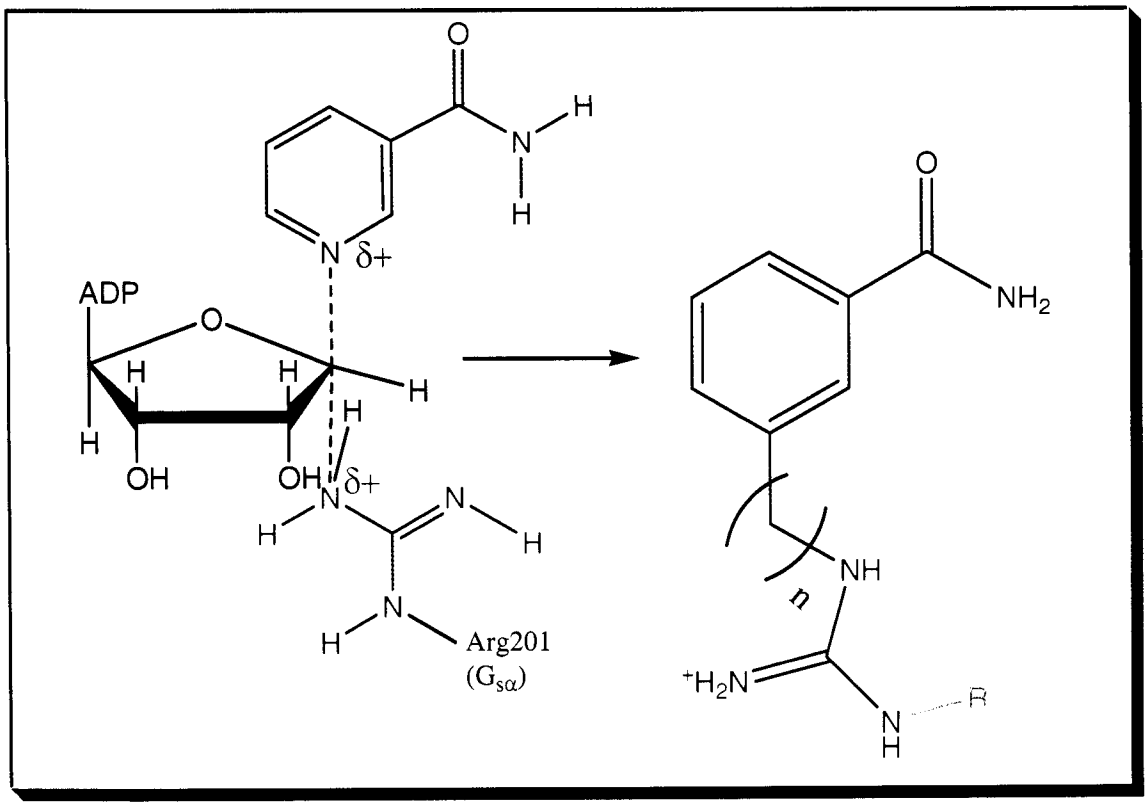


Figure 2.1. Design of bisubstrate analogue targeting CTA. Since two substrates are involved in the reaction, i.e. NAD and Arg201 of $G_{s\alpha}$, a bisubstrate analogue was designed which incorporates both the nicotinamide moiety (in red) and the arginine moieties (in blue). An R group on the other side of guanidine will make an extra contribution to the compound's affinity.

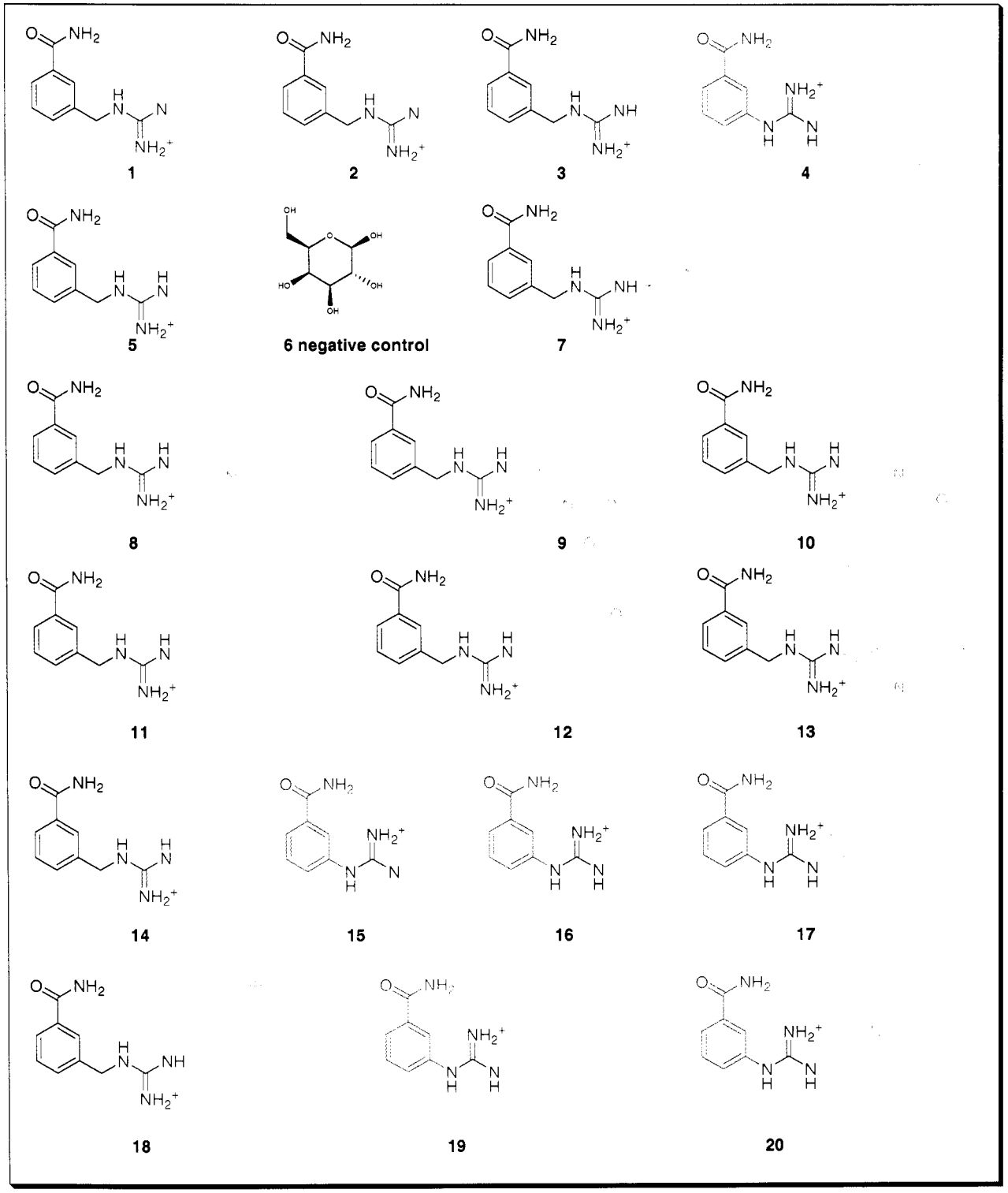


Figure 2.2. Structures of the first generation of inhibitors.

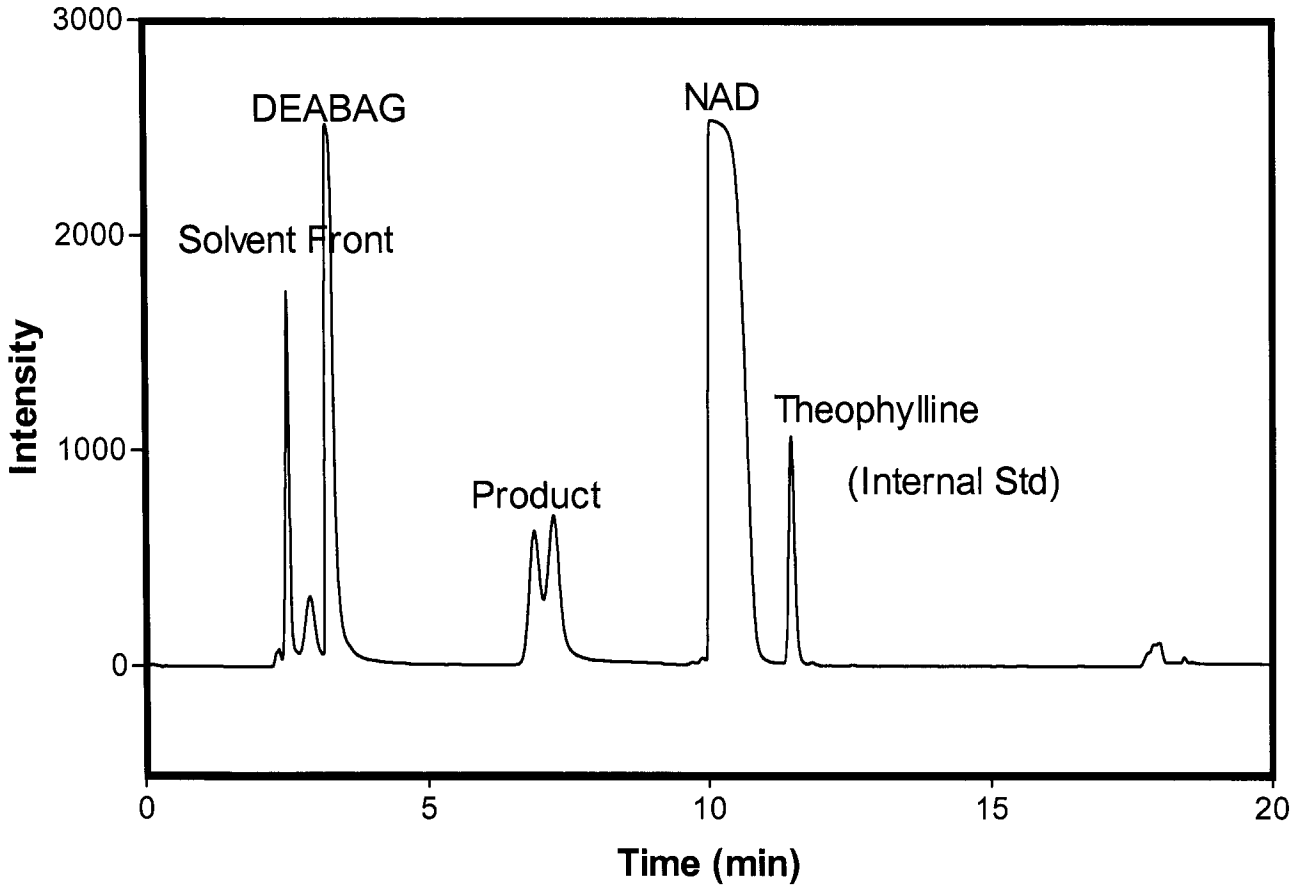


Figure 2.3. A typical chromatogram for an HPLC-based assay of CT. Each peak has been labeled. The two peaks overlapped in the product area have both been proved by mass spect and ^{1}H NMR to be the reaction product. As a result, they were both monitored.

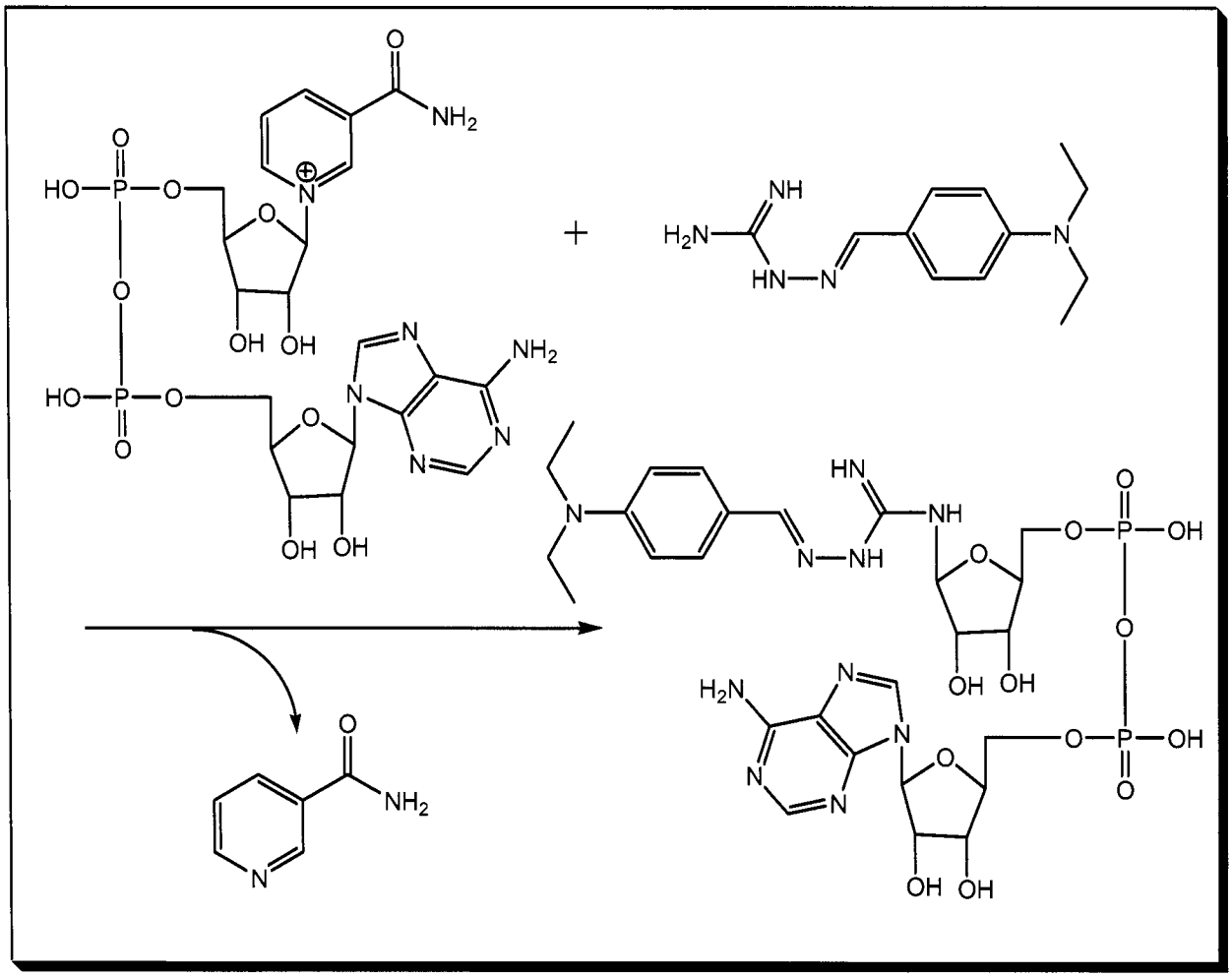


Figure 2.4. ADP-ribosylation reaction catalyzed by CT. In our kinetic study, we used DEABAG, a synthetic guandinium derivative, as an alternative ADP-ribose acceptor.

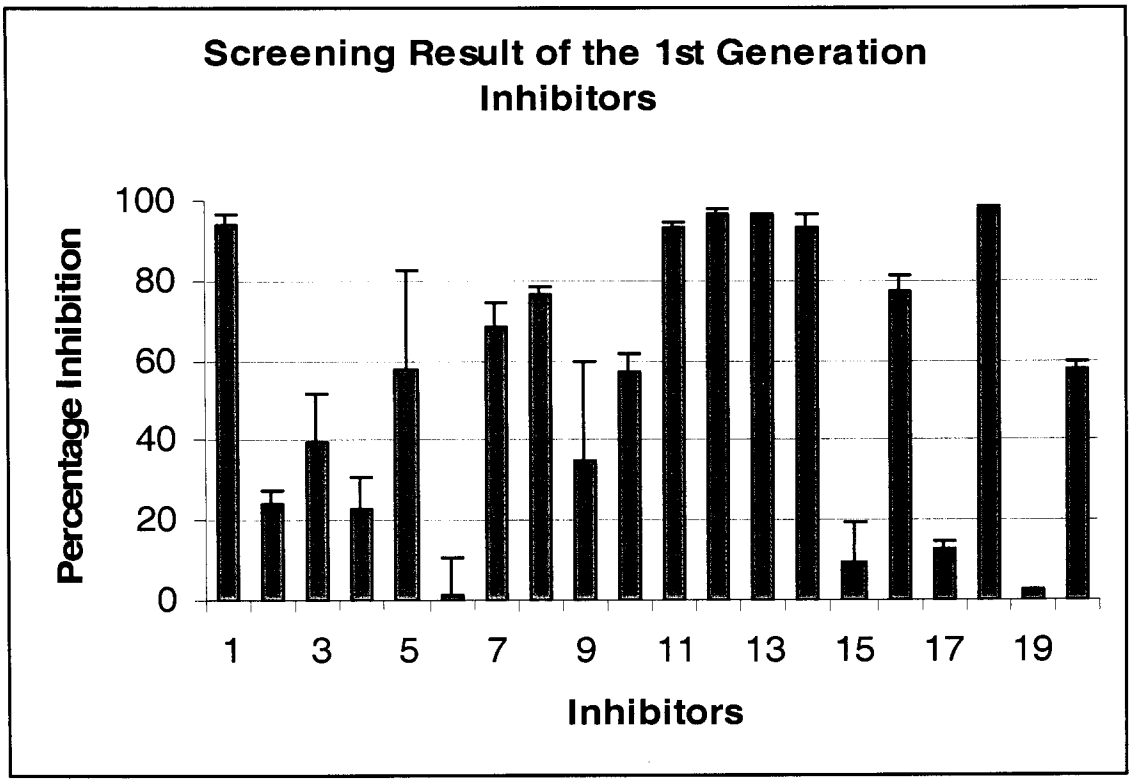


Figure 2.5. Screening results for the first generation of inhibitors in buffer PBS at a concentration of 1.0 mM. Compounds **1**, **11**, **13**, **14**, **18** demonstrate an inhibition of over 90%.

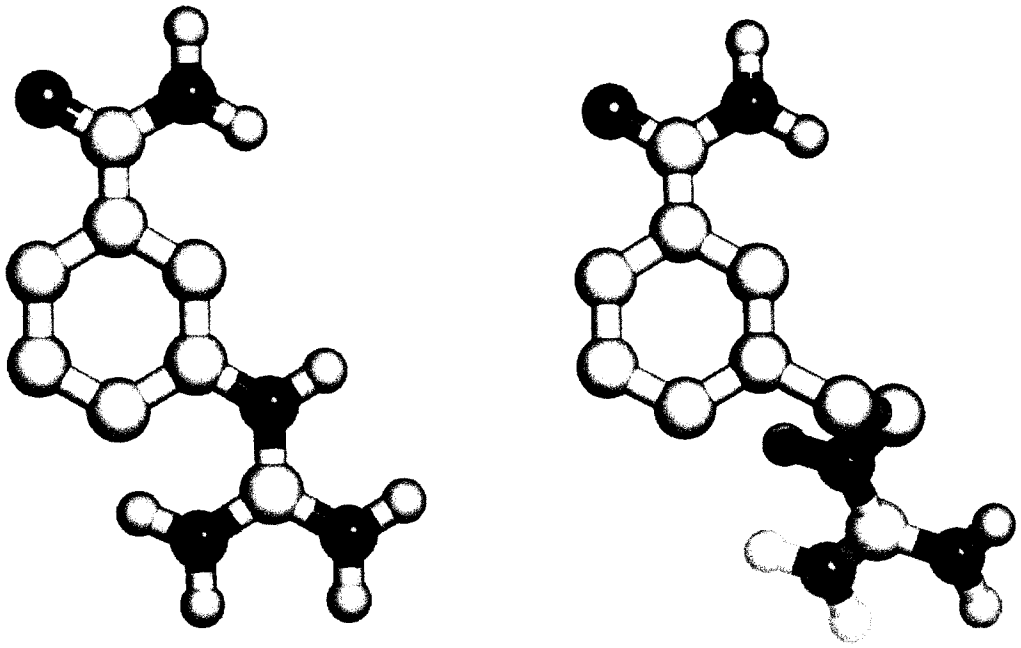


Figure 2.6. Comparison of CT inhibitors with different linkers between the nicotinamide analogue and guanidine moiety. As shown above, when there is no linker at all, the molecule forms a huge conjugated system and adopts a planar conformation (left). Whereas for inhibitors with a two-carbon linker, the molecule is more flexible (right). The conformation shown above is the low energy conformation from energy minimization with FLO96.

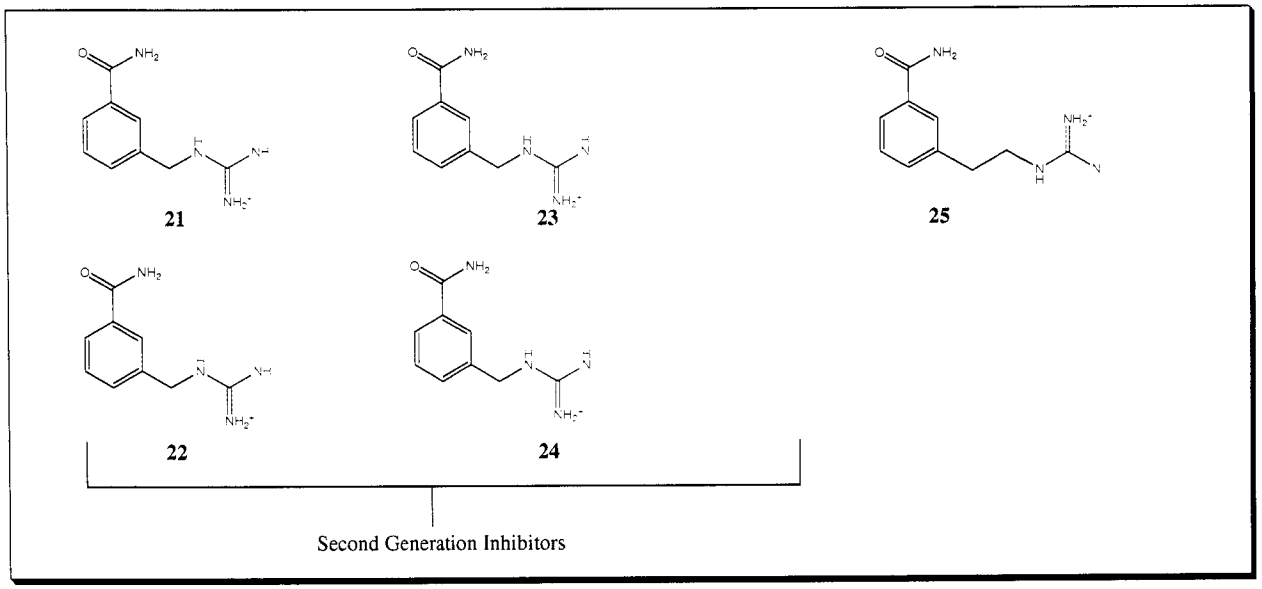


Figure 2.7. Structures of second generation inhibitors. A generic structure for inhibitors with a 2-carbon linker are also shown (**25**).

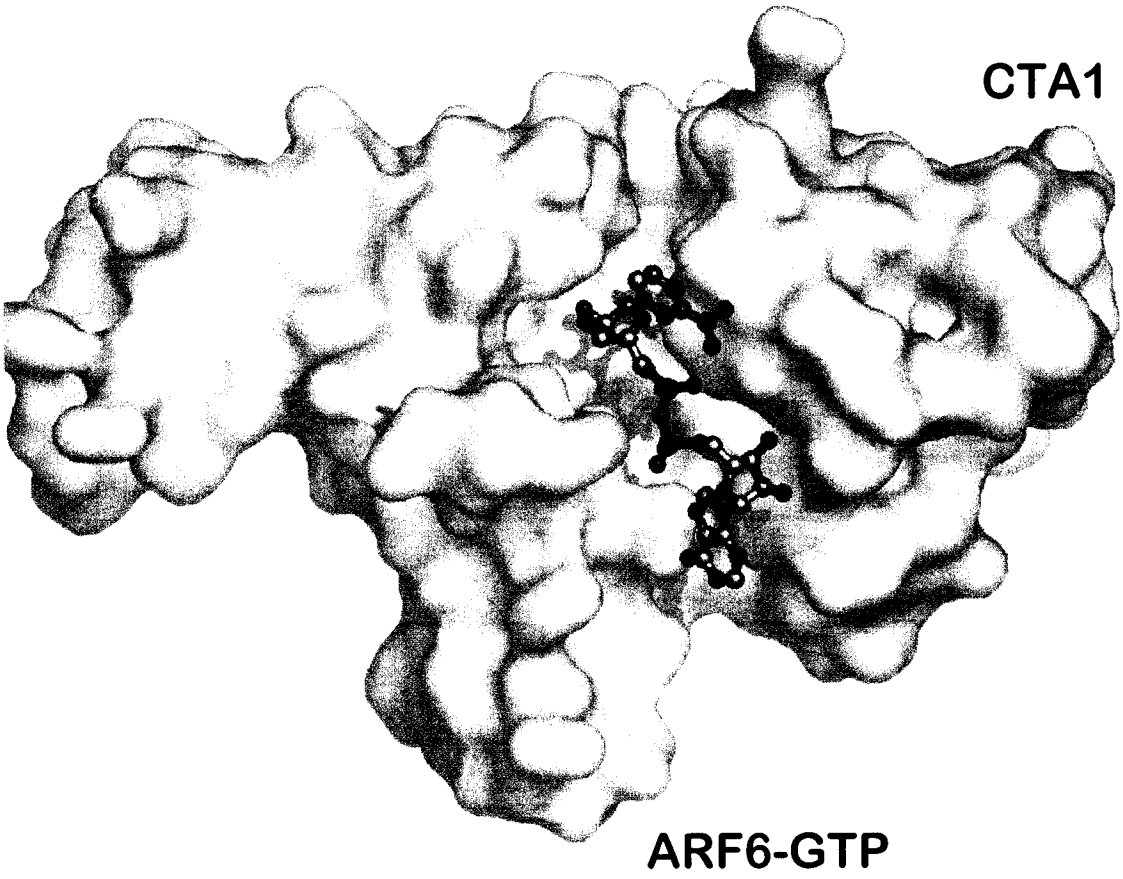


Figure 2.8. Surface representation of the NAD^+-occupied CTA1 active site (top view). The large active-site cleft is open to solvent from the top and both ends. The NAD^+ molecule binds snugly into the cleft. ARF6-GTP binds on the opposite face of CTA1 from NAD^+. This structure provides a solid foundation for molecular modeling. NAD molecule was color coded: carbon atoms are in orange, nitrogen atoms in blue, phosphorus atoms in pink and oxygen atoms in red. The surface representation is slightly transparent to allow the observation of the nicotinamide portion of NAD. This figure was produced by Pymol using PDB file 2A5F.

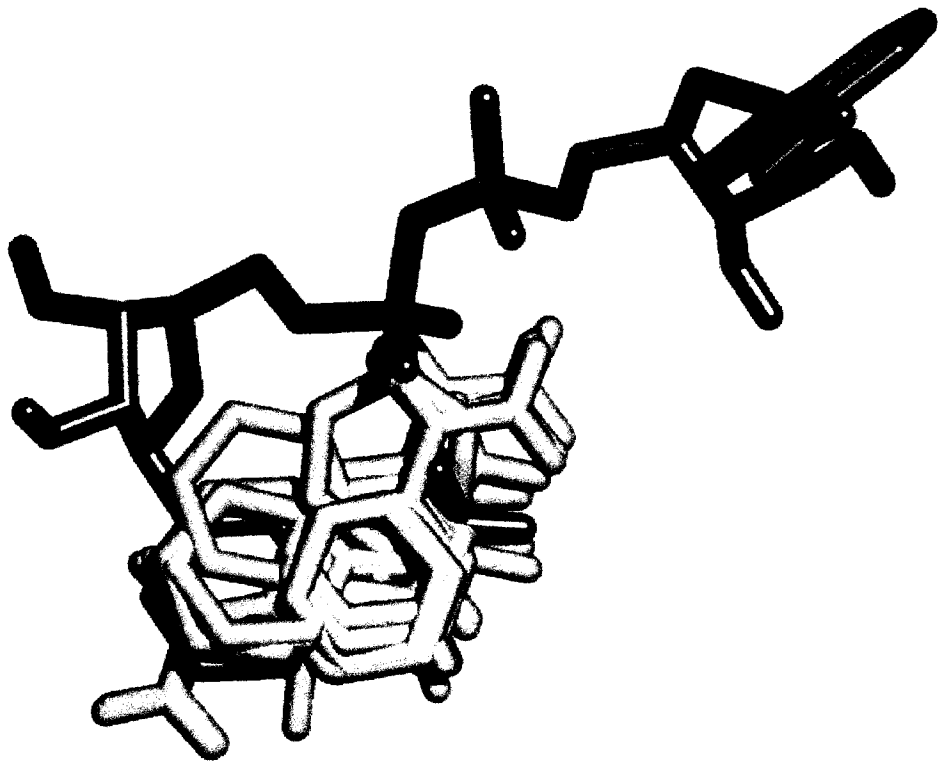


Figure 2.9. Docking results of benzamide into the NAD binding pocket of CTY30S. The results showed most of the low energy conformers bound to the same pocket as the nicotinamide portion of NAD, confirming that incorporating this moiety into the inhibitors is favorable. For clearance purposes, the surfaces have been omitted. Benzamide conformers are in green and NAD molecule is in violet.

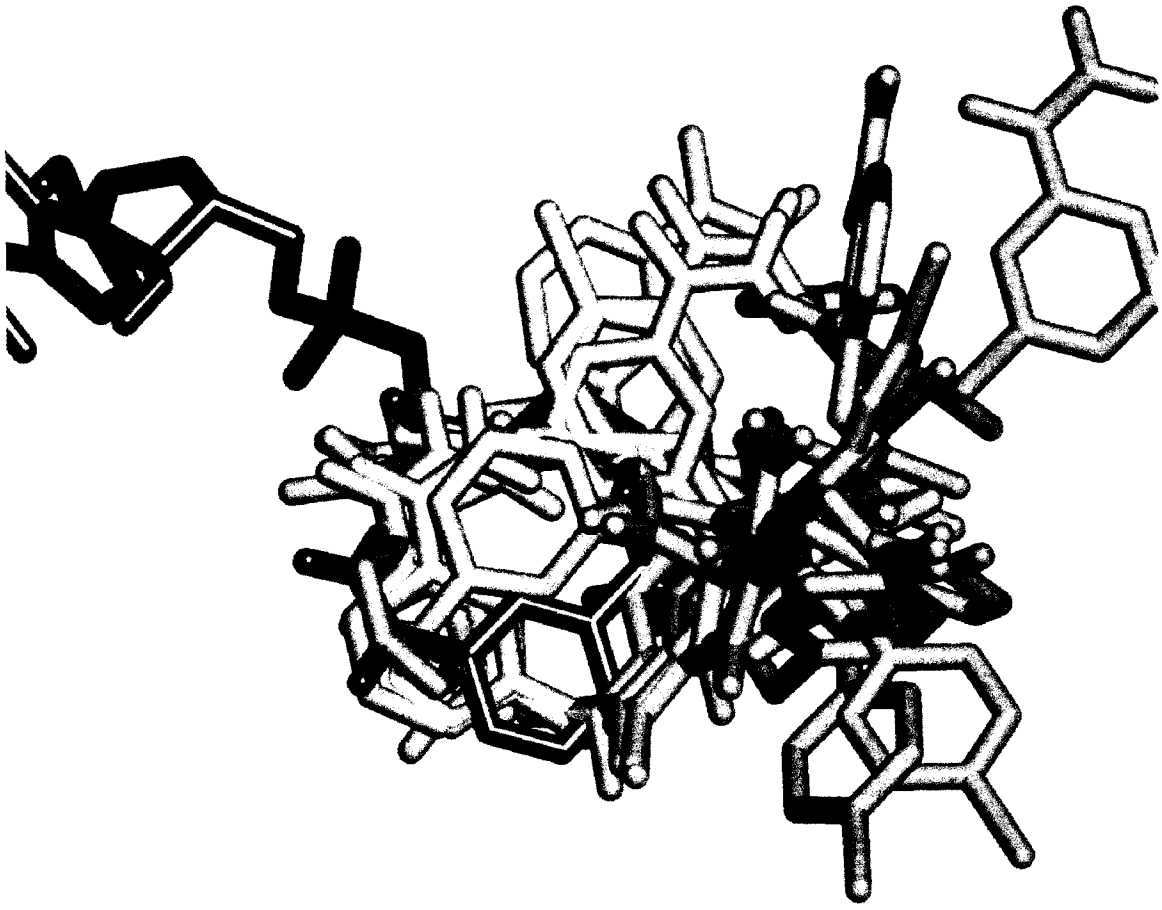
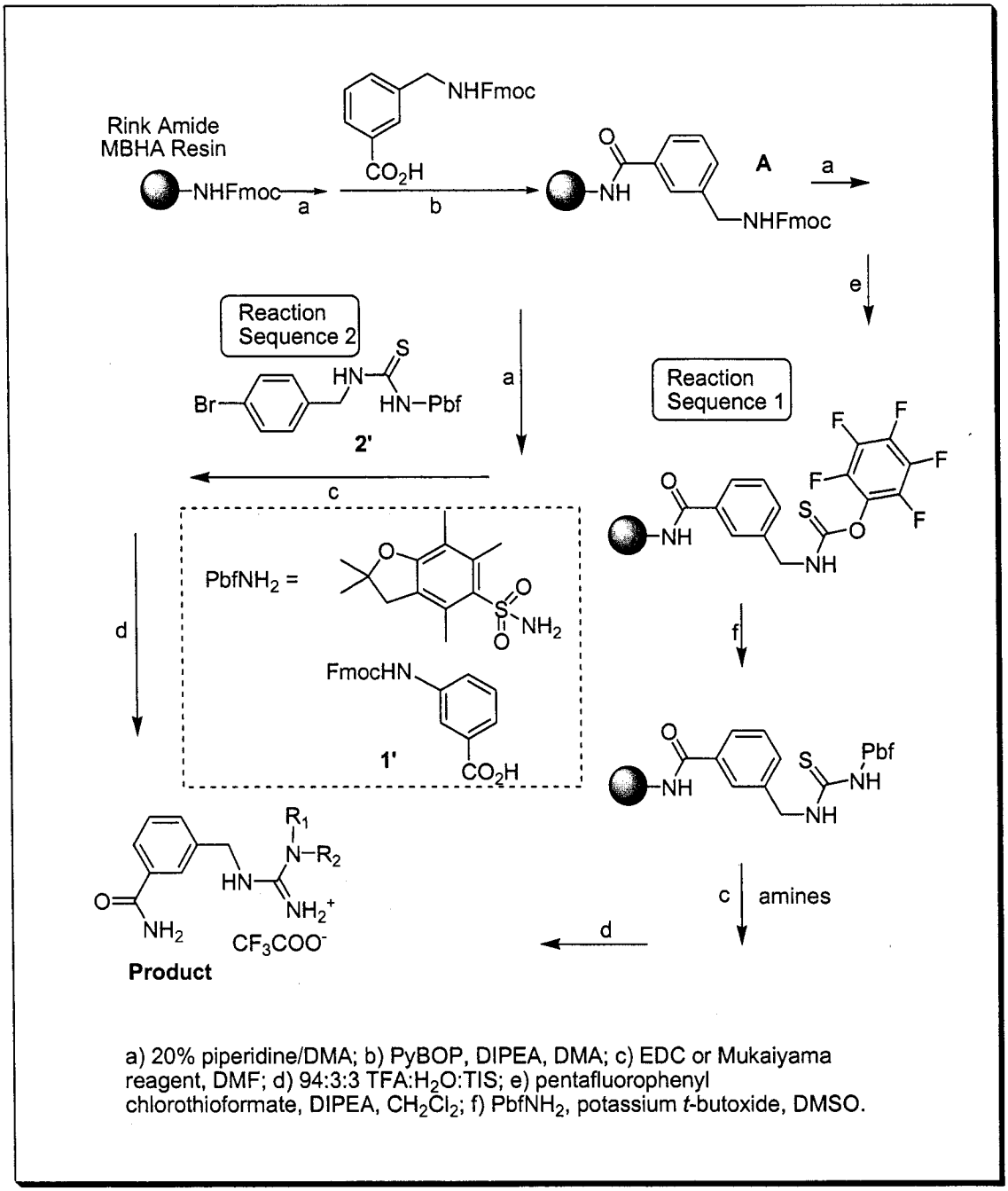
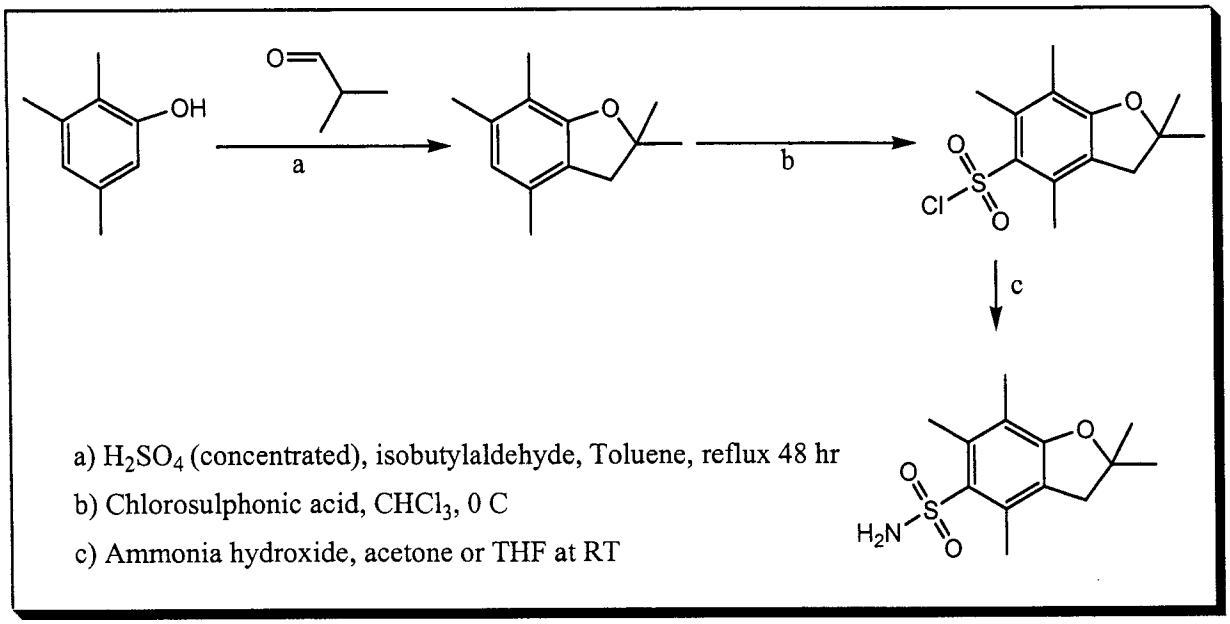


Figure 2.10. Docking results of compound **18**. There are some minor changes in the binding mode as suggested by our model. The guanidino group formed hydrogen bonds with Glu110 and Glu112 (residues not shown), pulling the molecule slightly out of the binding pocket.



a) 20% piperidine/DMA; b) PyBOP, DIPEA, DMA; c) EDC or Mukaiyama reagent, DMF; d) 94:3:3 TFA:H_2O:TIS; e) pentafluorophenyl chlorothioformate, DIPEA, CH_2Cl_2; f) $PbfNH_2$, potassium *t*-butoxide, DMSO.

Scheme 2.1. Solid-phase synthetic route of CT inhibitors. Both sequence 1 and 2 are applicable.



Scheme 2.2. Synthesis of precursor 2,2,4,6,7-pentamethylhydrobenzofuran-5-sulfonyl through a solution phase synthetic route.

Table 2.1. Summary of the screening results of the first generation inhibitors. R1 and R2 are substitutions on guanidine; n is the number of carbons in the linker. From this table, you can also get some information of structure-activity relationship. a. Samples screened at 1.0 mM concentration; b. Samples screened at 200 μM. Percentage inhibitions are only comparible in under the same screening conditions.

Compound	R1	R2	n	% Inhibition
1	Isopropyl	Isopropyl	1	94.22[a]
2	Ethyl	Ethyl	1	24.22[a]
3	Tert-butyl	H	1	39.42[a]
4	4-bromobenzyl	H	0	22.68[a]
5	Phenyl	H	1	57.80[a]
6	-	-	-	1.31[a]
7	3-(1-benzylpyrrilidine)	H	1	68.28[a]
8	N-ethyl-3-methyl-N-ethylbenzenamine	H	1	76.33[a]
9	4-(ethyl piperidine-1-carboxylate)	H	1	34.86[a]
10	4-ethylmorpholine	H	1	57.06[a]
11	Ethylphenyl	H	1	93.37[a]
12	4-methoxybenzyl	H	1	97.77[a]
13	4-benzylpiperidine	H	1	96.47[a]
14	Benzyl	H	1	93.12[a]
15	Isopropyl	Isopropyl	0	15.69[a]
16	Tert-butyl	H	0	77.35[a]
17	Phenyl	H	0	12.49[a]
18	4-bromobenzyl	H	1	98.31[a]
19	Benzyl	H	0	9.87[a]
20	4-benzylpiperidine	H	0	58.00[a]
21	2-Bromobenzyl	H	1	47.10[b]
22	3-Bromobenzyl	H	1	52.60[b]
23	1-(4-phenylbenzyl)	H	1	27.30[b]
24	2-methylnaphthalene	H	1	18.22[b]

Table 2.2. Cocrystallization and soaking results using artificial bi-substrate analogues. Soaking experiments and data collections were performed by C. J. O'Neal from the Hol lab (University of Washington). a. Form 2 crystal; b. Form 1 crystal. c. IC50 measured by kinetics studies. d. IC50 estimated based on screening.

Compound	IC50	Data Sets Collected	Result
1	125 μM[c]	5 mM soak for 5 min[a]	Active-site loop not displaced
3	2.3 mM[d]	5 mM soak for 3 min[a]	Active-site loop not displaced
11	270 μM[c]	10.9 mM cocrystal[b]	Active-site loop not displaced
14	200 μM[c]	5.9 mM cocrystal[a]	Active-site loop not displaced
15	350 μM[d]	5 mM soak for 6 min[a]	Active-site loop not displaced
17	4.2 mM[d]	5 mM soak for 6 min[a]	Active-site loop not displaced
18	44 μM[c]	10 mM cocrystal[b]	Active-site loop not displaced

Notes to Chapter 2

1. Zhou G., Parikn S. L., Tyler P. C., Evans G. B., Furneaux R. H., Zubkova O. V., Benjes P. A., Schramm V. L. Inhibitors of ADP-ribosylating bacterial toxins based on oxacarbenium ion character at their transitions states. *J. Am. Chem. Soc.* **2004**, 126, 5690-5698.

2. van den Akker F., Merritt E. A., Pizza M., Domenighini M., Rappuoli R., Hol W. G. The Arg7Lys mutant of heat-labile enterotoxin exhibits great flexibility of active site loop 47-56 of the A subunit. *Biochemistry* **1995**, 34, 10996-11004.

3. Gill D. M. Involvement of nicotinamide adenine dinucleotide in the action of cholera toxin *in vitro*. *Proc. Nal. Acad. Sci. USA* **1975**, 72, 2064-2068.

4. Osborne J. C. Jr., Stanley S. J., Moss J. Kinetic mechanisms of two NAD: arginine ADP-ribosyltransferases: the soluable, salt-stimulated transferase from turkey erythrocytes and chleragen, a toxin from *Vibrio cholerae*. *Biochemistry* **1985**, 5235-5240.

5. Li J., Zhang G., Zhang Z., Fan E. TFA-sensitive arylsulfonylthiourea-assisted synthesis of N,N'-substituted guanidines. *J. Org. Chem.* **2003,** 68, 1611-1614.

6. Manimala J. C., Anslyn E. V. Solid-phase synthesis of guanidinium derivatives from thiourea functionalities. *Eur. J. Org. Chem.* **2002**, 3909-3922.

7. Orner B. P., Hamilton, A. D. The guanidinium group in molecular recognition: design and synthetic approaches. *J. Inclusion Phenom. Macrocyclic Chem.* **2001**, 41, 141-147.

8. Narayanan J., Hartman P. A., and Graves D. J. Assay of heat-labile enterotoxins by their ADP-ribosyltransferase activities. *J. Clin. Microbiol.* **1989**, 27, 2414-2419.

9. Soman G., Narayanan J., Martin B. L., Graves D. J. Use of substituted (benzylidineamino)guanidines in the study of guanidine group specific ADP-ribosyltrasferase. *Biochemistry* **1986**, 25, 4113-4119.

10. Loesberg C., van Rooij H., Smets L. A. Meta-iodobenzylguanidine (MIBG), a novel high-affinity substrate for cholera toxin that interferes with cellular mono(ADP-ribosylation). *Biochim. Biophys. Acta* **1990**, 1037, 92-99.

11. Fersht A. A guide to enzyme catalysis and protein folding. W. H. Freeman and Company, New York, **1999**.

12. Larew J. S., Peterson J. E., Graves D. J. Determination of the kinetic mechanism of arginine-specific ADP-ribosyltransferases using a high performance liquid chromatographic assay. *J. Biol. Chem.* **1991**, 266, 52-57.

13. Galloway T. S., Van Heyningen S. Binding of NAD+ by cholera toxin. *Biochemistry* **1987**, 244, 225-230.

14. Moss J., Manganiello V. C., Vaughan M. Hydrolysis of nicotinamide adenine dinucleotide by choleragen and its A protomer: possible role in the activation of adenylate cyclase. *Proc. Natl. Acad. Sci. USA* **1976**, 73, 4424-4427.

15. Mekalanos J. J., Collier R. J., Romig W. R. Enzymic activity of cholera toxin. I. New method of assay and the mechanism of ADP-ribosyl transfer. *J. Biol. Chem.* **1979**, 254, 5849-5854.

16. O'Neal C. J., Amaya E. I., Jobling M. G., Holmes R. K., Hol W. G. Crystal structures of an intrinsically active cholera toxin mutant yield insight into the toxin activation mechanism. *Biochemistry* **2004**, 43, 3772-3782.

17. O'Neal C. J., Jobling M. G., Holmes R. K., Hol W. G. Structural basis for the activation of cholera toxin by human ARF6-GTP. *Science* **2005**, 309, 1093-1096.

18. Seidler J., McGovern S. L., Doman T. N., Shoichet B. K. Identification and prediction of promiscuous aggregating inhibitors among known drugs. *J. Med. Chem.* **2003**, 46, 4477-4486.

19. Teng Z., Moore J. L., Craig A. R., Hickey G., Schelp C., Wei T. Reduction of non - specific binding in assays. *U.S. Pat. Appl. Publ.* US 2006105472, **2006**.

20. Falcovitz-Gerassi, Y., Tsinberg P. Alleviation of non -specific binding in microarray assays. *U.S. Pat. Appl. Publ.* US 2006040407, **2006**.

21. Peterfi Z., Kocsis B. Comparison of blocking agents for an ELISA for LPS. *J. Immunoassay* **2000**, 21, 341-354.

22. Papavassiliou A. G., Bohmann D. Optimization of the signal-to-noise ratio in south-western assays by using lipid-free BSA as blocking reagent. Nucleic Acids Res. **1992**, 20, 4365-4366.

23. Brorson S. H. Bovine serum albumin (BSA) as a reagent against non-specific immunogold labeling on LR-White and epoxy resin. *Micron* **1997**, 28, 189-195.

24. Kononenko V. L., Shimkus J. K., Giddings J. C., Myers M. N. Feasibility studies on photophoretic effects in field-flow fractionation of particles. *J. Liq. Chromatogr. Relat. Technol.* **1997**, 20, 2907-2929.

25. Bowen W. P., Jerman J. C. Nonlinear regression using spreadsheets. *Trends Pharmacol. Sci.* **1995**, 16, 413-417.

26. McMartin C., Bohacek R. S. QXP: Powerful, rapid computer algrithms for structure-based drug design. *J. Comput. Aided Mol. Design* **1997**, 11, 333-344.

27. ThistleSoft, Inc. FLO96, ThistleSoft, Colebrook, CT, **1996**.

Chapter 3

Background and Introduction of Strategies for Targeting Protein-protein Interactions

Introduction

Many proteins, including enzymes, exert their biological activity partially if not completely by interacting with other proteins. To modulate the activity of such proteins by blocking these interactions is a very active research area. Protein-protein interactions play a pivotal role in many key biological processes, such as viral self assembly, signal transduction and cell proliferation, growth, differentiation and programmed death.[1-5] Therefore, they have become a very important class of targets for developing cellular probes and therapeutic reagents. A range of biologically significant protein interactions have been investigated to design potent modulators. Some of the successes have been reviewed previously.[6-11] With the explosion in genomics and proteomics, an enormous amount of new targets for therapeutics based on protein–protein interactions are rapidly emerging. The growth in structural information on these interactions, in combination with advanced computational and combinatorial synthesis methods, provides an excellent starting point for structure-based inhibitor design. In spite of the remarkable successes in small-molecule inhibitor design targeting definite enzyme active sites, the development of small-molecule modulator of protein-protein interactions remains a formidable challenge. As large families of enzymes bind to the same substrate, it is possible to use the knowledge gained and the compound libraries that have been precedently synthesized to target a single member. Afterwards, we can quickly design compounds that target new

members of the family at a much faster pace.[12-14] This, of course, significant accelerates the drug discovery process. In the case of protein–protein interactions, even if similarities are observed between some interfaces, it is quite unusual that binding sites are preserved amongst protein interfaces. This hurdle can be mainly attributed to the following aspects possessed by protein–protein interactions: firstly, the connection between proteinaceous ligands and the design of small molecules is often indirect and inconspicuous. The key residues that are involved are usually indefinite; secondly, when a protein–protein interaction is formed, it usually has a large buried surface, typically over 1500 $Å^2$, which is a serious challenge for small molecules to address; a third difficulty is that the binding interface of the protein partners is often not contiguous, which makes it especially difficult for simple synthetic molecules to mimic the interaction; another common feature for the protein–protein interfaces is that they are usually lack of special features. A selectivity issue arises as a result; finally, there are a very limited number of small molecules identified to disrupt the protein–protein interactions from high-throughput screening.[15] On the other hand, since there is only a limited number of direct biological assays to accurately detect enzymatic activity,[16] it makes it even more challenging to address protein–protein interactions.

Facing such a challenging task, a variety of design approaches has been developed, for example, cross-linked interfacial peptides,[17-22] helix mimetics,[23-29] and small-molecule leads [30-37] from various drug candidate screening programs etc. The future holds great promise for custom-made, potent inhibitors of protein–protein interactions on demand.

Various Protein–Protein Interfaces

Amongst the large variety of protein–protein interactions in living organisms, a wealth of proteins form either transient or permanent complex to exert their biological functions.[38] Since there are such diverse interfaces, it is very important to survey the druggable properties of the target interface before starting a drug discovery process.

Protein complexes are formed from identical subunits (homo-oligomers) or from different subunits (hetero-oligomers).[39] These oligomers can be formed directly during protein synthesis or on the encounter of different subunits. Protein complexes also have various stabilities. Permanent complexes are very stable, and their subunits remain assembled upon association, however, targeting the interface of a permanent oligomer is rather difficult and thus is not the mainstream of existing research; whereas other complexes exist only transiently and sometimes in a dynamic equilibrium. Their chain associates are more easily broken. This means that compounds may be identified which, upon binding to the contact surface of one subunit, prevent interaction with the other subunit in nonpermanent oligomers. Furthermore, the formation of some complexes depends on the presence of an effector molecule, or changes in ambient physiological conditions. In this project, I will focus on the inhibitor design targeting a nonpermanent interface.

In addition, the shape of the interface is another important parameter to consider for drug discovery. It is more difficult to obtain potent inhibitors for those with a flat interface other than those with a well-defined pocket.[40] The hetero-oligomers tend to have planer surfaces than homodimers. This suggests that an interface incorporating a nonpermanent homodimer will be the most attractive target for drug discovery. Besides

the cavities that are present at an interface, shape complementarity is also of great importance. It will be more difficult to generate potent competitive inhibitors for those chains which are tightly packed. It has a higher probability to design a potent inhibitor for an interface that is loosely packed, where there are not a lot of direct interactions, than that is tightly packed. They will offer more opportunities to improve affinity. Water molecules will fill in to satisfy H-bond requirements if the complementarities between interfaces are deficient. The displacement of key bound water molecules by the inhibitor might enhance its affinity because of a favorable entropic effect.

Key Features of Protein–Protein Associations

Although there are daunting challenges concerning inhibitor design that targets protein–protein interfaces, some key features of the interface hold new promises for this intricate task. First of all, some "hot spots" exist in the binding interface. A milestone work by Clarkson and Wells reveals that the binding energy of human growth hormone (hGH) and its first-bound receptor (hGHbp) can be simply ascribed to a small set of interfacial residuals.[41] This result suggests that it is possible to mimic an important domain of a partner protein with a relatively small molecule. Further research disclosed that a small fraction of the interfacial residues accounted for the majority of the total binding energy, whilst other residues did not contribute much if at all. These residues constituted a key binding region, the so-called "hot spots". Thermal dynamic features such as enthalpy, entropy and heat capacity also effect in the formation of a stable complex. The association constant K, which describes the affinity of a ligand for a protein that binds to it, is the most important parameter. It is determined by ΔG between

the associated and unassociated states.[42-44] Even though there are some contradicting opinions about the correlations between ΔH or ΔS and the values of ΔG, it has been generally accepted that hydrophobic interactions provide the key driving force for protein–protein complex formation. Based on the work by Cochran et al., the driving force for disruption could derive from either enthalpy-favored protein stabilization or entropy-favored ligand exchange.[45] A third feature of the protein–protein interface is the cooperativity of the interactions. A common misconception is that protein-protein interactions rely solely on the properties of the interacting surface. The fact is they also depend in part on how these interactions modify the internal structure of the protein. In the event of a positive cooperativity, the protein structure is tightened (**Scheme 3.1**). In the meantime, its affinity is expected to increase. In a negatively cooperative binding event, the protein structure is loosened and its affinity is decreased as a result.[46-49]

Strategies for Protein Surface Recognition

Each protein surface is unique in terms of charge, hydrophobicity, and hydrophilicity. These protein surfaces are important for many biological functions. Therefore, synthetic molecules that satisfy both the electrostatic (including van der Waals, hydrogen bond and salt bridge) and structural features of the protein surfaces are desired to obstruct the protein-protein interactions. These molecules could enable the regulation of a range of biological processes. To date, strategies toward protein surface recognition fall into two categories: biological strategies and chemical strategies. The biological strategies include mainly artificial antibodies, [50,51] miniature proteins,[52-54] functional nucleotides,[55] and unnatural biopolymers.[56] The chemical strategies encompass cross-

linked interfacial peptides and other peptidomimetics, helix mimetics, and small-molecule leads from library screening (**Scheme 3.2**), among which the helix mimetics approach seems to be the most intriguing one and it has attracted people's attention in the past few years. Hamilton and coworkers exploited a terphenyl scaffold to mimic α-helices. Subsequent work published by the same group further modified and optimized it to a terephthalate-based system, which presents the key residues of an α-helix at *i*, *i*+4, and *i*+7 positions in a very similar fashion.[23-29] Another extensively studied approach for α-helix mimicry is cross-linked interfacial peptides. People have proposed to link the main chain or the side chain of the peptides to constrain the peptides to fold in an α-helical conformation.[17-22] In this project, I am going to employ both the aforementioned two approaches to advance the development of potent inhibitor to disrupt protein–protein interactions.

Some Typical Systems with a Helix-Helix Interface

As diversified as the protein-protein interfaces are, there are quite a few systems with helix-helix interfaces which people are actively pursing (**Table 3.1**). Bcl–xL is the dominant regulator of apoptosis.[57, 58] On the other hand, Bak accelerates programmed cell death by binding to and antagonizing the repressor Bcl-2. A Bcl–xL–Bak complex structure determined by NMR spectroscopy shows that a helical region of Bak (residues 72–87) binds to a hydrophobic cleft on the surface of Bcl-xL. Furthermore, the crucial residues for binding were shown by alanine scanning to be Val74, Leu78, Ile81, and Ile85, which is commensurate with the *i*, *i*+4, *i*+7 and *i*+11 positions along one face of the α-helix. Another broadly investigated system is the MDM2-p53 system.[59-61] Murine

double-minute 2 (MDM2) is a regulator of a tumor suppressor p53, which protects cells by responding to genotoxic stresses with an induction of cell cycle arrest or apoptosis. Because downregulation of p53 by overexpression of MDM2 has been implicated in many human sarcomas, it has drawn a growing interest to exploit the p53-MDM2 interface as a chemotherapeutic target. MDM2 recognizes the three conserved hydrophobic residues (Phe19, Trp23, and Leu26) within a 15-residue N-terminal region of the p53. The side chains of the key p53 residues form a hydrophobic binding motif on one face of a short α-helix, which fits tightly to the hydrophobic cleft of MDM2. Tremendous amount of efforts has been made to get efficacious small-molecule inhibitor targeting this system.[30-37]

Aside from the two well-known systems, some other systems with helix-helix interfaces spring up lately. MyoA, an unusual motor in malaria invasion machinery to the host cell, is in complex with the MyoA-tail Interacting Protein (MTIP), inducing a major conformational change.[62] Three key hydrophobic MyoA residues, Leu804, Val807, and Ile811, aligning very well on one side of the tail helix, point toward the MTIP surface and fit snugly into the hydrophobic pockets. Since it has been reported that MyoA-tail peptide is able to block the parasite growth, this interface opens new avenues for anti-malarial drug design.

Recently, the crystal structure of dormancy survival regulator C-terminal domain ($DosR^c$) of tuberculosis was solved.[63] Detailed studies indicate two subunits A and B have a helix-helix dimerization interface, engaging 15 residues including Thr198, Val202 and Thr205. The AB dimer is crucial for the ensuing DNA-binding event. Ultimately,

DosR activates the transcription of *Mycobacteria tuberculosis* (MTB) into a long-lasting dormant state.

Tuberculosis

Tuberculosis (TB) is a devastating infectious disease caused by pathogenic *Mycobacteria tuberculosis*, which primarily affects the lungs. It has become a global public health epidemic and the death toll is nearly 2 million each year (**Figure 3.1**).[64] Its ability to perdure in the host tissues for an extended period of time makes it a most perilous disease in the world. At present, one third of the world population is under the threat of latent TB (LTB).[64] Under proper conditions, latent TB will be reactivated and turn into active pulmonary TB. Moreover, with the surge of HIV, it has been reported that one third of the HIV patients worldwide are co-infected with TB, accounting for 15 million infection cases.[65] TB spreads by close contact with patients who have active pulmonary TB. A patient with active TB can infect up to 20 people. Upon initial infection, only 40% of individuals develop primary active TB. The rest of the infection will become latent TB.[66] The risk of reactivation is highly dependent on the infected person's immune status. Currently, treatment for TB is the most effective when it is in its active form, curing 95% of active TB patients.[67,68] Nevertheless, drug resistant TB strains are virtually untreatable.[69,70] Latent TB is more difficult to of treat because the bacilli are in an obstinate state and requires prolonged chemotherapy.

Invasion of *Mycobacteria Tuberculosis*

TB disperses by close contact with patients who have active pulmonary TB. The aerosol droplets containing live *M. tuberculosis* are produced during coughing, sneezing, or talking. They are inhaled and enter the pulmonary alveoli of the lung. *M. tuberculosis* is then ingested by alveolar macrophages where the bacteria are either directly killed or grow to a limited extend inside phagosome. The bacillus replicates slowly but continuously inside macrophages. Another fate for *M. tuberculosis* is that it is taken up by dendritic cells and then transported to local lymph nodes for the recruitment and activation of additional macrophages. Infection with *M. tuberculosis* does not necessarily transform into disease. About 2 to 6 weeks after initial infection, cell-mediated immunity develops in most infected people. Several cytokines such as interleukins-12 (IL-12) are produced by infected alveolar macrophages to stimulate a series of responses for the activation of additional macrophages to differentiate the *M. tuberculosis* infection. These specialized macrophages and T-lymphocytes surround infected macrophages to form caseous granulomas that limit further replication and spread of the organism and it usually results in the death of the vast majority of the bacteria.[71-74] A small proportion of surviving bacteria may remain dormant but alive inside caseous granulomas. This pathological state is termed as latent TB infection, which can sustain for a lifetime of an infected individual with a potential of resuming growth and causing active disease many years or decades after initial infection (**Figure 3.2**). Only 40% of individuals after initial infection actually develop primary active TB. The infection for the rest of people becomes latent TB, which can be reactivated to active TB over a span from months to several decades with 2-23% chance per lifetime.[66] When the latently infected individuals

suffer from immune system depletion by HIV, immunosuppressive drugs, aging, starvation etc., *M. tuberculosis* will be reactivated and could be disseminated to other organs.

Dormancy Survival Regulator and Tuberculosis

Albeit the cause of latent TB is still unclear, the onset of latent TB is usually associated with reduced oxygen tension (hypoxia) and nitric oxide exposure.[75,76] The exact same set of 47 *M. tuberculosis* genes is quickly up-regulated in response to each of these stimuli. A two-component DosS-DosR system is induced.[77] It has been demonstrated that the sensor kinases DosS and DosT phosphorylate their conserved histidine and then transfer the phosphate to Asp54 of DosR.[78-80] The phosphrylation of Asp54 considerably enhances the binding affinity of DosR to its cognate DNA sequence (the DosR box).[78] DosR binds to the DNA duplex at a promoter region of the hypoxic response gene *acr* (also known as *hspX*).[81] Mutagenesis studies confirm that binding site residuals are strictly conserved and are required for DNA binding activity. Computer analysis also identified a consensus sequence recognized by DosR. A variant of some kind is located upstream of almost all *M. tuberculosis* genes that are rapidly induced by hypoxia.[81,82]

DosR is believed to mediate the transition of *M. tuberculosis* into dormancy and may contribute to latency. Disruption of DosR was reported to increase the virulence of *M. tuberculosis* in mice [83] but leads to attenuation of *M. tuberculosis* virulence in guinea pigs.[84] Therefore, DosR could be a good target for drug development towards latent tuberculosis.

Current Treatment for Tuberculosis

Several tests are available for TB diagnosis varying by sensitivity and accuracy. The tuberculosis skin test (PPD, Mantoux) is the most commonly used test for early detection of latent TB with a few restrictions.[81] It uses purified protein derivative (PPD) of *M. tuberculosis* to test the immune response. The Sputum test applies to patients with active TB, but its relatively low accuracy (40-50%) is a big drawback. Other tests that offer greater accuracy take up to six weeks and culture is usually required.

Long term treatment with a combination of drugs is required to cure TB. This makes the treatment an intimidating challenge to the global public health. The standard "short" course treatment for tuberculosis (TB), if it is active, is isoniazid, rifampicin, pyrazinamide and ethambutol for two months, then isoniazid and rifampicin alone for a further four months. The patient is considered cured at six months. For latent tuberculosis, the standard treatment is six to nine months of isoniazid alone. Daily administration of isoniazid is effective in lowering the risk of TB reactivation for latent TB patient by 60-90% over a period of 6-12 months.[67,68]

The Structure of Dormancy Survival Regulator C-terminal Domain ($DosR^C$)

The crystal structure of DosR C-terminal domain was determined at 2.0 Å resolution. The asymmetric unit contains eight molecules, assembled into two $DosR^C$ tetramers. The $DosR^C$ subunit comprises four α-helices designated as α7, α8, α9 and α10. In each tetramer, $DosR^C$ forms two equivalent dimers, AB and CD (**Figure 3.3**). The interfaces between AB and CD are essentially identical with a two-fold axis between two subunits. A third interface puts AB and CD together with a two-fold axis between B

and C subunits. This two-fold is perpendicular to, but non-intersecting with, the two dimer dyads, generating a non-crystallographic two-fold screw axis perpendicular to all three two-folds in the tetramer.

The α10 dimer interface. The dimerization interface between subunits A and B involves mainly the α10 helix plus additional residues from the α7-α8 loop. A pseudo-two-fold symmetry axis is observed in each dimer. This interface buries 1,000 $Å^2$ of solvent-accessible surface (**Figure 3.4a**). Some other amino acids are also involved in the interactions. Thr198 contributes the most to the interface, accounting for about 17% of the total buried surface area, while Gly164, Arg196, Val202, and Thr205, each contribute 10-14%. Thr198, Val202, and Thr205 (TVT) are the key residues responsible for the contacts between the two α10 helices in the AB dimer and they align very well on the same side of the α-helix. The TVT makes a combined contribution of 40% of the entire α10-α10 surface.[63]

The α7/α8 tetramer interface. The second type of interface occurs between subunits B and C, organizing dimers AB and CD into a tetrameric unit. B and C are related by a non-crystallographic two-fold axis, into a tetramer. The dimer-dimer contacts involve mainly residues from helices α7 and α8 plus the α8-α9 loop. This interface has a buried surface of approximately 1,500 $Å^2$. The α7/α8 tetramer interface is significantly larger than the AB dimer interface (**Figure 3.4b**). There are several direct and water-mediated hydrogen bonds and numerous van der Waals contacts. Arg173 and Phe175 contribute the most to the interface, ~12% and 24%, respectively. The side chain of Phe175 in each subunit protrudes out and sticks into a hydrophobic pocket created by the

apolar portion of the sidechains of Arg155, Leu158, Gly159, Ala204, Leu207, and Lys208 (plus Leu147 in subunit B or Pro213 in subunit C) of the interacting subunit. This hydrophobic cavity is substantial, and without which the sidechain of Phe175 would extend into the solvent.[63]

Scope of This Project

The research work described here in this project is largely based on the structure-based drug design methodology. The purpose of the project is to find superb α-helix mimetics to probe the protein-protein interaction and thereby achieve the goal of modulating the key biological processes at will. The structural features of the α10/α10 interface of DosR provide a desirable system to investigate protein-protein interaction between a homodimer. Strategies developed in this project could also afford solutions to other similar systems listed earlier. The first half of Chapter 4 describes the design, modeling, synthesis and evaluation of small-molecule α-helix mimetics. The compounds are terphenyls in essence, with substitutions at specific positions. A virtual screening using the program FLO96 predicted the low energy binding modes and explained inhibitory results from a gel-shift binding assay. It is worth noting that the crystallographic work and part of the initial biological assays were carried out by our collaborator in Professor Wim G. J. Hol's laboratory at the Biomolecular Structural Center in University of Washington.

The second half of chapter 4 focuses on the elucidation of peptide α-helix. This part of work was initiated by the optimization of the truncated native α10 sequence. The most potent linear peptide was further truncated to a minimum of eight residues to fully represent "TVT" on the same side of the α-helix. Computer modeling results suggest that side chain cyclization could enhance α-helicity to a greater extent. The percentage of α-helical content was characterized by Circular Dichrism (CD) and two-dimensional

NMR (ROESY and TOCSY). Long linear peptides with unnatural amino acids were also synthesized and evaluated.

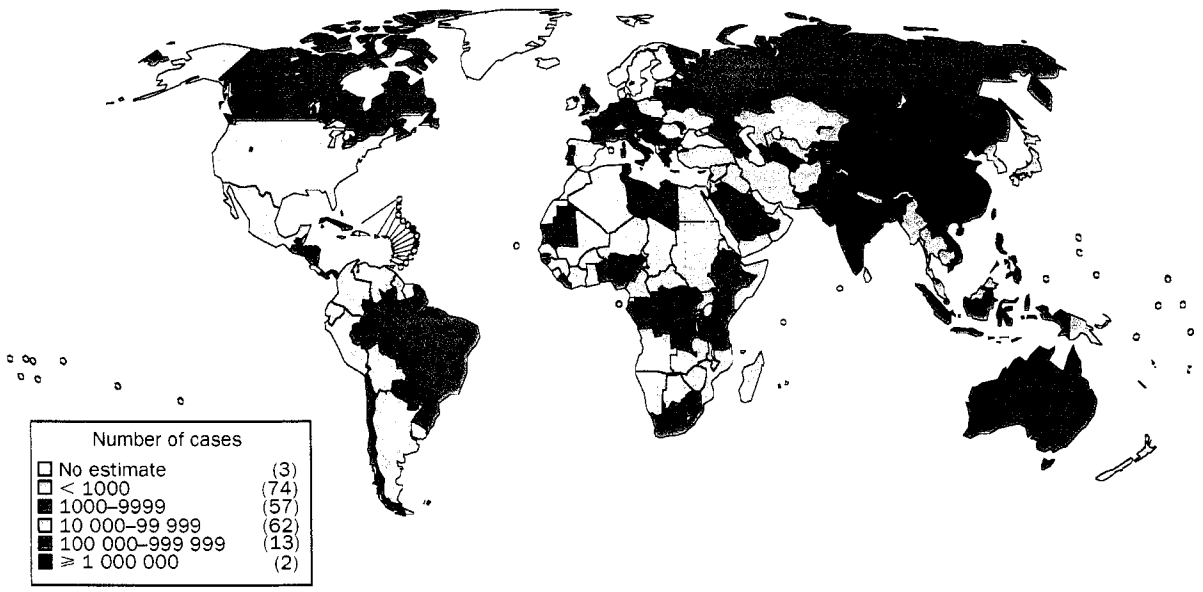
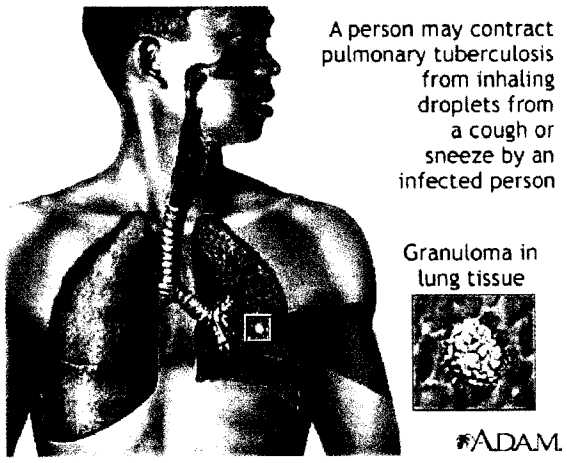


Figure 3.1. Estimated number of new tuberculosis cases in the year 2001. This figure was adapted from ref. 68 without further modification.



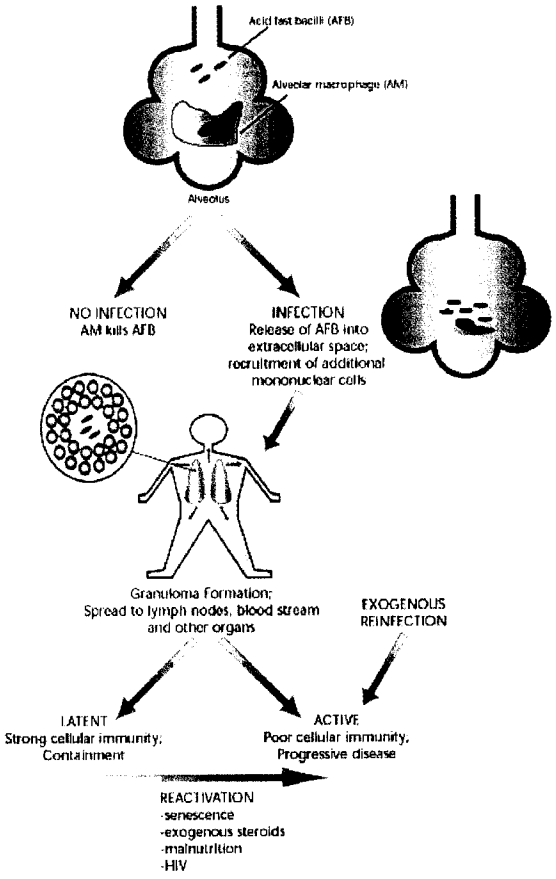


Figure 3.2. Tuberculosis infection. The bacteria contained in aerosol droplets are inhaled and enter the lung of a person. *M. tuberculosis* is then ingested by alveolar macrophages. The bacillus replicates slowly but continuously inside macrophages. After that additional macrophages are activated. Infection with *M. tuberculosis* does not necessarily transform into disease. It could turn into latent TB. When the infected individual's immune system is going down, MTB is reactivated and can develop into active TB again. This figure was adapted partially from the website of Research Center Borstel.

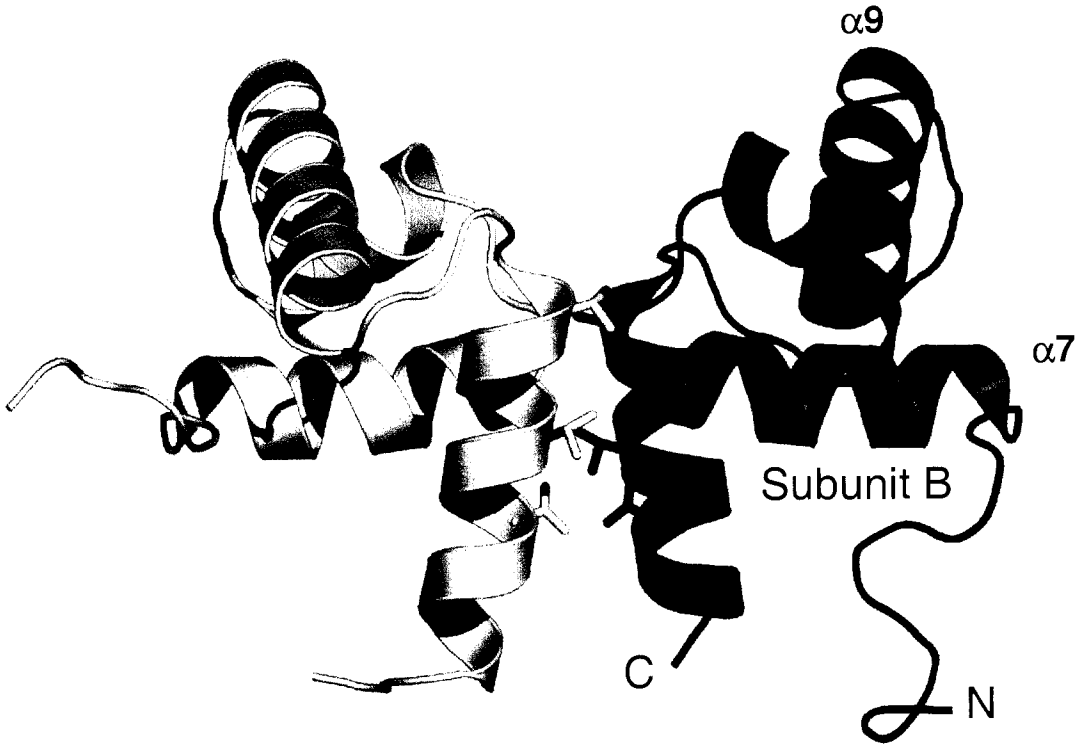


Figure 3.3. Crystal structure of the AB dimer of $DosR^C$. It is a homodimer and each monomer comprises four α-helices—α7, α8, α9, α10. "C" stands for the C-terminal, "N" for N-terminal. Three important residues T198, V202, T205 are presented on the same side of the α10 helix and they contribute to over 40% of the total interaction. This figure was kindly provided by Dr. Goragot Wisechaisri from Prof. Hol's lab.

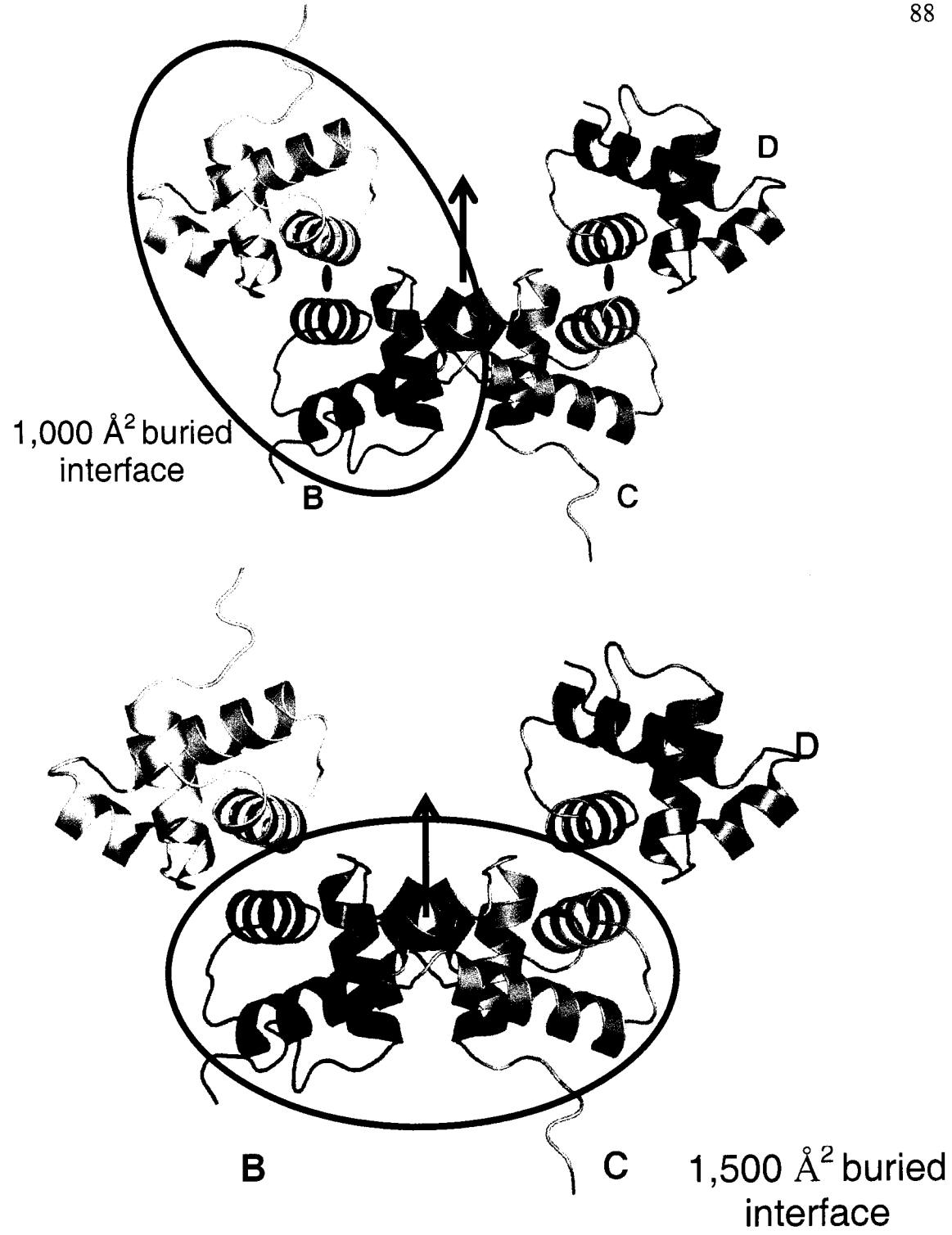
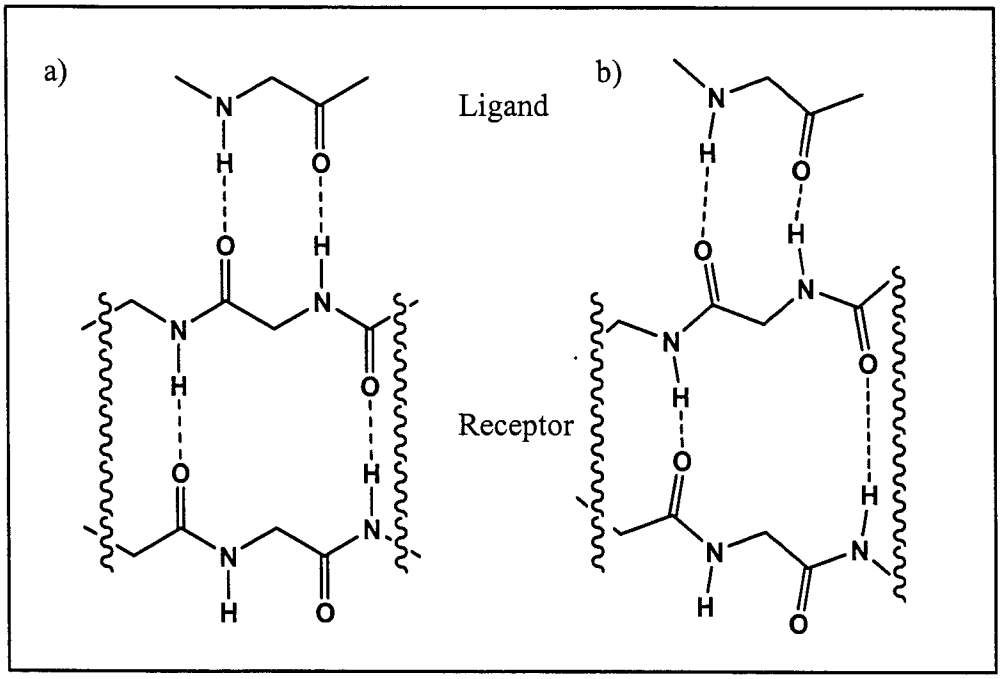
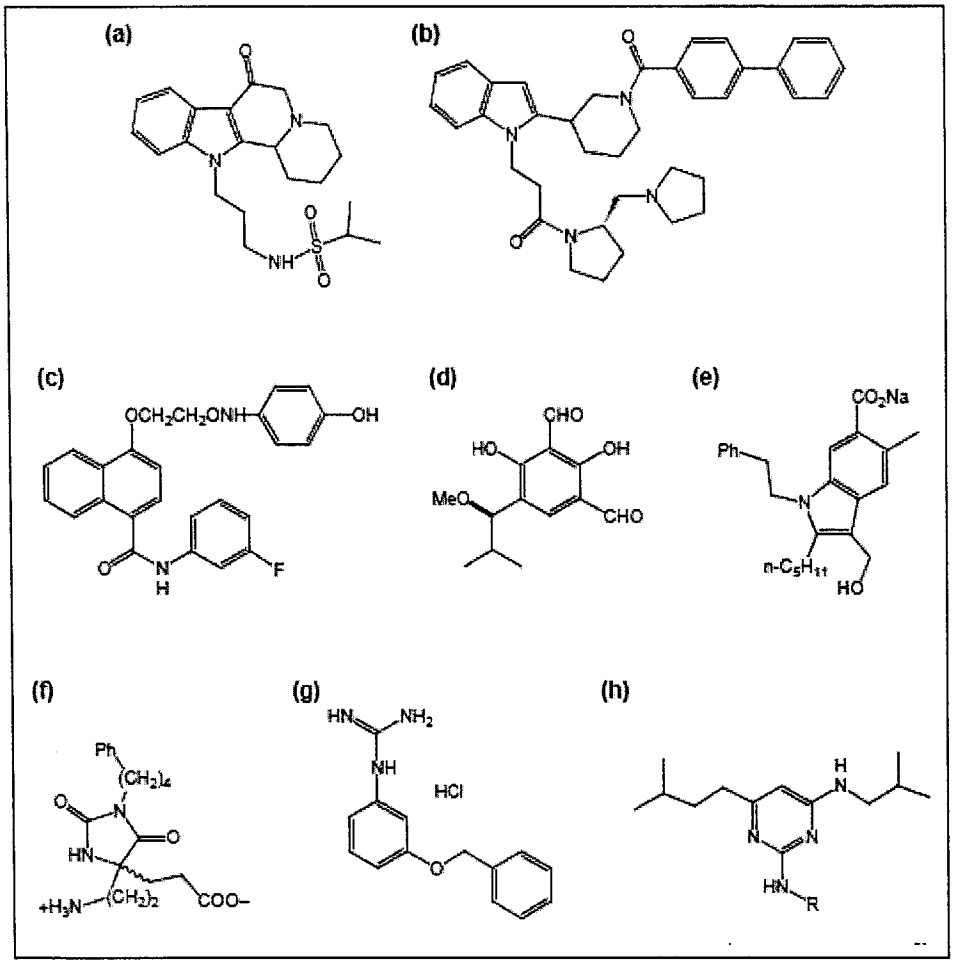


Figure 3.4. The crystal structure of DosR^C tetramer. a) The AB interface has a buried surface of ~1000 Å^2; b) The BC interface has a buried surface of ~ 1500 Å^2. The corresponding surfaces have been circled by red ovals. This figure was kindly provided by Dr. Goragot Wisechaisri from Prof. Hol's lab.



Scheme 3.1. Structural models for a) positive and b) negative cooperativity; the lower pair of hydrogen bonds is within a protein receptor and the upper pair of hydrogen bonds is formed upon the binding of a peptide ligand. This figure was reproduced based on the concept of outlined in ref. 13.



Scheme 3.2. Summary of recently reported small-molecule inhibitors of protein–protein interactions: (a) ZipA–FtsZ,[30] (b) ZipA–FtsZ,[31] (c) HOX–PBX,[32] (d) Sam68–Fyn SH3,[33] (e) MAGI3–PTEN,[34] (f) insulin–receptor,[35] (g) CD4–gp120,[36] (h) nuclear receptor–coactivator.[37]

Table 3.1. Some typical systems that incorporate a protein–protein interface. Their interface properties and drug development status are also listed for comparison.

Entry	Protein Interface	Dimer Type	Helix-helix Interface	Inhibitor Status
1	DosR–DosR	Homodimer	Yes	No.
2	Bcl–xL–Bak	Heterodimer	Yes	Well developed
3	MDM2–p53	Heterodimer	Yes	Well developed
4	MyoA–MTIP	Heterodimer	Yes	Peptides only
5	CD4-gp120	Heterodimer	No	Well developed

Notes to Chapter 3

1. Pellegrini M., Haynor D., Johnson J. M. Protein interaction networks. *Expert Rev. Proteomics* **2004**, 1, 239-249.

2. Jordan J. D., Landau E. M., Iyengar R. Signaling networks: the origins of cellular multitasking. *Cell* **2000**, 103, 193-200.

3. Hipfner D. R., Cohen S. M. Connecting proliferation and apoptosis in development and disease. *Nat. Rev. Mol. Cell Biol.* **2004**, 5, 805-815.

4. Oliner J. D., Pietenpol J. A., Thiagalingam S., Gvuris J., Kinzler K. W., Vogelstein B. Oncoprotein MDM2 conceals the activation domain of tumour suppressor p53. *Nature* **1993**, 362, 857-860.

5. Cordon-cardo C., Latres E., Drobnjak M., Oliva M. R., Pollack D., Woodruff J. M., Marechal V., Chen J. D., Brennan M. F., Levine A. J. Molecular abnormalities of mdm2 and p53 genes in adult soft tissue sarcomas. *Cancer Res.* **1994**, 54, 794-799.

6. Toogood P. L. Inhibition of protein-protein association by small molecules: approaches and progress. *J. Med. Chem.* **2002**, 45, 1543-1558.

7. Berg T. Modulation of protein-protein interactions with small organic molecules. *Angew. Chem. Int. Ed. Engl.* **2003**, 42, 2462-2481.

8. Gadek T. R., Nicholas J. B. Small molecule antagonists of proteins. *Biochem. Pharmacol.* **2003**, 65, 1-8.

9. Arkin M. R., Wells J. A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* **2004**, 3, 301-317.

10. Pagliaro L., Felding J., Audouze K., Nielsen S. J., Terry R. B., Krog-Jensen C., Butcher S. Emerging classes of protein-protein interaction inhibitors and new tools for their development. *Curr. Opin. Chem. Biol.* **2004**, 8, 442-449.

11. Zhao L., Chmielewski J. Inhibiting protein-protein interactions using designed molecules. *Curr. Opin. Struct. Biol.* **2005**, 1, 31-34.

12. Chene P. Drugs targeting protein-protein interactions. *ChemMedChem.* **2006**, 4, 400-411.

13. Chene P. Inhibition of the p53-MDM2 interactions: Targeting a protein–protein interface. *Mol. Cancer Res.* **2004**, 2, 20-28.

14. Chene P. p53 as a drug target in cancer therapy. *Expert Opin. Ther. Patents* **2001**, 11, 923-935.

15. Yin H., Hamilton A. D. Strategies for targeting protein-protein interactions with synthetic agents. *Angew. Chem. Int. Ed. Engl.* **2005**, 27, 4130-4163.

16. Lakey J. H., Raggett E. M. Measuring protein–protein interactions. *Curr. Opin. Struct. Biol.* **1998**, 8, 119-123.

17. Li P., Roller P. P. Cyclization strategies in peptide derived drug design. *Curr. Top. Med. Chem.* **2002**, 2, 325-341.

18. Sarabia F., Chammaa S., Ruiz A. S., Ortiz L. M., Herrera F. J. L. Chemistry and biology of cyclic peptides of medicinal and biological interest. *Curr. Med. Chem.* **2004**, 11, 1309-1332.

19. Taylor J. W. The synthesis and study of side-chain lactam-bridged peptides. *Biopolymers.* **2002**, 1, 49-75.

20. Taylor J. W., Jin Q. K., Sbacchi M., Wang L., Belfiore P., Garnier M., Kazantzis A., Kapurniotu A., Zaratin P. F., Scheideler M. A. Side-chain lactam-bridge conformational constraints differentiate the activities of salmon and human calcitonins and reveal a new design concept for potent calcitonin analogues. *J. Med. Chem.* **2002**, 5, 1108-1121.

21. Tian Y., Ramesh C. V., Ma X., Naqvi S., Patel T., Cenizal T., Tiscione M., Diaz K., Crea T., Arnold E., Arnold G. F., Taylor J. W. Structure-affinity relationships in the gp41 ELDKWA epitope for the HIV-1 neutralizing monoclonal antibody 2F5: effects of side-chain and backbone modifications and conformational constraints. *J. Pept. Res.* **2002**, 6, 264-76.

22. Jackson D. Y., King D. S., Chmielewski J., Singh S., Schultz P. G. General approach to the synthesis of short .alpha.-helical peptides. *J. Am. Chem. Soc.* **1991**, 24, 9391-9392.

23. Kutzki O., Park H. S., Ernst J. T., Orner B. P., Yin H., Hamilton A. D. Development of a potent Bcl-x(L) antagonist based on alpha-helix mimicry. *J. Am. Chem. Soc.* **2002**, 40, 11838-11839.

24. Ernst J. T., Becerril J., Park H. S., Yin H., Hamilton A. D. Design and application of an alpha-helix-mimetic scaffold based on an oligoamide-foldamer strategy: antagonism of the Bak BH3/Bcl-xL complex. *Angew. Chem. Int. Ed. Engl.* **2003**, 5, 535-539.

25. Ernst J. T., Kutzki O., Debnath A. K., Jiang S., Lu H., Hamilton A. D. Design of a protein surface antagonist based on alpha-helix mimicry: inhibition of gp41 assembly and viral fusion. *Angew. Chem. Int. Ed. Engl.* **2002**, 41, 278-281.

26. Yin H., Hamilton A. D. Terephthalamide derivatives as mimetics of the helical region of Bak peptide target Bcl-xL protein. *Bioorg. Med. Chem. Lett.* **2004**, 6, 1375-1379.

27. Yin H., Lee G. I., Park H. S., Payne G. A., Rodriguez J. M., Sebti S. M., Hamilton A. D. Terphenyl-based helical mimetics that disrupt the p53/HDM2 interaction. *Angew. Chem. Int. Ed. Engl.* **2005**, 18, 2704-2707.

28. Yin H., Lee G. I., Sedey K. A., Rodriguez J. M., Wang H. G., Sebti S. M., Hamilton A. D. Terephthalamide derivatives as mimetics of helical peptides: disruption of the Bcl-x(L)/Bak interaction. *J. Am. Chem. Soc.* **2005**, 15, 5463-5468.

29. Kim I. C., Hamilton A. D. Diphenylindane-Based Proteomimetics Reproduce the Projection of the *i*, *i*+3, *i*+4, and *i*+7 Residues on an α-Helix. *Org. Lett.* **2006**, 8, 1751-1754.

30. Jennings L. D., Foreman K. W., Rush T. S., Tsao D. H. H., Mosyak L., Li Y. H., Sukhdeo M. N., Ding W. D., Dushin E. G., Kenny C. H. Design and synthesis of indolo[2,3-alpha]quinolizin-7-one inhibitors of the ZipA-FtsZ interaction. *Bioorg. Med. Chem. Lett.* **2004**, 14, 1427-1431.

31. Jennings L. D., Foreman K. W., Rush T. S., Tsao D. H. H., Mosyak L., Kincaid S. L., Sukhdeo M. N., Sutherland A.G., Ding W. D., Kenny C. H. Combinatorial synthesis of substituted 3-(2-indolyl)piperidines and 2-phenyl indoles as inhibitors of Zip A-FtsZ interaction. *Bioorg. Med. Chem.* **2004**, 12, 5115-5131.

32. Ji T., Lee M., Pruitt S. C., Hangauer D. G. Privileged scaffolds for blocking protein-protein interactions: 1,4-disubstituted naphthalene antagonists of transcription factor complex HOX-PBX/DNA. *Bioorg. Med. Chem. Lett.* **2004**, 14, 3875-3879.

33. Oneyama C., Agatsuma T., Kanda Y., Nakano H., Sharma S. V., Nakano S., Narazaki F., Tatsuta K. Synthetic inhibitors of proline-rich ligand-mediated protein-protein interaction: potent analogs of UCS15A. *Chem. Biol.* **2003**, 10, 443-451.

34. Fujii N., Haresco J. J., Novak K. A. P., Stokoe D., Kuntz I. D., Guy R. K. A selective irreversible inhibitor targeting a PDZ protein interaction domain. *J. Am. Chem. Soc.* **2003**, 125, 12074-12075.

35. Tan C., Wei L. H., Ottensmeyer F. P., Goldfine I., Maddux B. A., Yip C. C., Batey R. A., Kotra L. P. Structure-based de novo design of ligands using a three-dimensional model of the insulin receptor. *Bioorg. Med. Chem. Lett.* **2004**, 14, 1407-1410.

36. Boussard C., Klimkait T., Mahmood N., Pritchard M., Gilbert I. H. Design, synthesis and evaluation of potential inhibitors of HIV gp120-CD4 interactions. *Bioorg. Med. Chem. Lett.* **2004**, 14:2673-2676.

37. Rodriguez A. L., Tamrazi A., Collins M. L., Katzenellenbogen J. A. Design, synthesis, and in vitro biological evaluation of small molecule inhibitors of estrogen receptor a coactivator binding. *J. Med. Chem.* **2004**, 47, 600-611.

38. Li S., Armstrong C. M., Bertin N., Ge H., Milstein S., Boxem M., Vidalain P. O., Han J. D., Chesnau A., Hao T., Goldberg D. S., Li N., Martinez M., Rual J. F., Lamesch P., Xu L., Tewari M., Wong S. L., Zhang L. V., Berriz G. F., Jacotot L., Vaglio P., Reboul J., Hirozane-Kishikawa T., Li Q., Gabel W., Elewa A., Baumgartner B., Rose D. J., Yu H., Bosak S., Sequerra R., Fraser A., Mango S. E., Saxton W. M., Strome S., Van Den Heuvel S., Piano F., Vandenhaute J., Sardet C., Gerstein M., Doucette-Stamm L., Gunsalus K. C., Harper J. W., Cusick E., Roth F. P., Hill D. E., Vidal M. A map of the interactome network of the metazoan C. elegans. *Science* **2004**, 303, 540-543.

39. Nooren I. M. A., Thornton J. M. Diversity of protein-protein interactions. *EMBO J.* **2003**, 22, 3486-3492.

40. Jones S., Thornton J. M. Searching for functional sites in protein structures. *Proc. Natl. Acad. Sci. USA* **1996**, 93, 13-20.

41. Clackson T., Wells J. A. A hot spot of binding energy in a hormone-receptor interface. *Science* **1995**, 267, 383-386.

42. Weber G. Thermodynamics of the association and the pressure dissociation of oligomeric proteins. *J. Phys. Chem.* **1993**, 97, 7108-7115.

43. Makhatadze G. I., Privalov P. L. Energetics of protein structure. *Adv. Protein Chem.* **1995**, 48, 307-425.

44. Kresheck G. C., Vitello L. B., Erman J. E. Calorimetric studies on the interaction of horse ferricytochrome c and yeast cytochrome c peroxidase. *Biochemistry* **1995**, 34, 8398-8405.

45. Cochran A. G. Antagonists of protein-protein interactions. *Chem. Biol.* **2000**, 7, 85-94.

46. Williams D. H., Maguire A. J., Tsuzuki W., Westwell M. S. An analysis of the origins of a cooperative binding energy of dimerization. *Science* **1998**, 280, 711-714.

47. Williams D. H., Stephens E., Zhou M. How can enzymes be so efficient? *Chem. Commun.* **2003**, 16, 1973-1976.

48. Williams D. H., Stephens E., Zhou M. Ligand binding energy and catalytic efficiency from improved packing within receptors and enzymes. *J. Mol. Biol.* **2003**, 329, 389-399.

49. Williams D. H., Calderone C. T., O_Brien D. P., Zerella R. Changes in motion vs. bonding in positively vs. negatively cooperative interactions. *Chem. Commun.* **2002**, 1266-1267.

50. Dall'Acqua W., Carter P. Antibody engineering. *Curr. Opin. Struct. Biol.* **1998**, 8, 443-450.

51. Vaughan T. J., Osbourn J. K., Tempest P. R. Human antibodies by design. *Nat. Biotechnol.* **1998**, 16, 535-539.

52. Martin F., Toniatti C., Salvati A. L., Venturini S., Ciliberto G., Cortese R., Sollazzo M. The affinity-selection of a minibody polypeptide inhibitor of human interleukin-6. *EMBO J.* **1994**, 13, 5303-5309.

53. Ku J., Schultz P. G. Alternate protein frameworks for molecular recognition. *Proc. Natl. Acad. Sci. USA* **1995**, 92, 6552-6556.

54. Choo Y., Klug A. Designing DNA-binding proteins on the surface of filamentous phage. *Curr. Opin. Biotechnol.* **1995**, 6, 431-436.

55. Green L. S., Jellinek D., Bell C., Beebe L. A., Feistner B. D., Gill S. C., Jucker F. M., Janjic N. Nuclease-resistant nucleic acid ligands to vascular permeability factor/vascular endothelial growth factor. *Chem. Biol.* **1995**, 2, 683-695.

56. Gellman S. H. Foldamers: A Manifesto. *Acc. Chem. Res.* **1998**, 31, 173-180.

57. Klasa R. J., Gillum A. M., Klem R. E., Frankel S. R. Oblimersen bcl-2 antisense: facilitating apoptosis in anticancer treatment. *Antisense Nucleic Acid Drug Dev.* **2002**, 12, 193-213.

58. Gutierrez-Puente Y., Zapata-Benavides P., Tari A. M., Lopez-Berestein G. Bcl-2-related antisense therapy. *Semin. Oncol.* **2002**, 29, 71-76.

59. Haupt Y., Maya R., Kazaz A., Oren M. Mdm2 promotes the rapid degradation of p53. *Nature* **1997**, 387, 296-299.

60. Honda R., Tanaka H., Yasuda H. Oncoprotein MDM2 is an ubiquitin ligase E3 for tumor suppressor p53. *FEBS Lett.* **1997**, 420, 25-27.

61. Kubbutat M. H. G., Jones S. N., Vousden K. H. Regulation of p53 stability by Mdm2. *Nature* **1997**, 387, 299-303.

62. Bosch J., Turley S., Daly T. M., Bogh S. M., Villasmil M. L., Roach C., Zhou N., Morrisey J. M., Vaidya A. B., Bergman L. W., Hol W. G. Structure of the MTIP-MyoA complex, a key component of the malaria parasite invasion motor. *Proc. Natl. Acad. Sci. USA* **2006**, 103, 4852-4857.

63. Wisedchaisri G., Wu M., Rice A. E., Roberts D. M., Sherman D. R., Hol W. G. Structures of Mycobacterium tuberculosis DosR and DosR-DNA complex involved in gene activation during adaptation to hypoxic latency. *J. Mol. Biol.* **2005**, 354, 630-641.

64. World Health Organization. Intensified action needed to halt a quarter of a million TB/HIV deaths a year to people living with AIDS. *Joint News Release WHO/International AIDS Society WHO/41.* 16 August, **2006**.

65. World Health Organization. Global tuberculosis control: surveillance, planning, financing. *WHO report 2005*, World Health Organization, Geneva, **2005**.

66. Parrish N. M., Dick J. D., Bishai W. R. Mechanisms of latency in *Mycobacterium tuberculosis*. *Trends Microbiol.* **1998**, 6, 107-112.

67. Frieden T. R., Sterline T. R., Munsiff S. S., Watt C. J., Dye C. Tuberculosis. *Lancet* **2003**, 362, 887-899.

68. Blumberg H. M., Leonard M. K. Jr., Jasmer R. M. Update on the treatment of tuberculosis and latent tuberculosis infection. *JAMA* **2005**, 293, 2776-2784.

69. World Health Organization: Concern over extensive drug resistant TB strains that are virtually untreatable. *News release by WHO.* 5 September, **2006**.

70. Altman K. L. Drug-resistant TB in South Africa draws attention from U. N. *New York Times* 6 September, **2006**.

71. Smith I. *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. *Clin. Microbiol. Rev.* **2003**, 16, 463-496.

72. Frieden T. R., Sterline T. R., Munsiff S. S., Watt C. J., Dye, C. Tuberculosis. *Lancet* **2003**, 362, 887-899.

73. Stewart G. R., Robertson B. D., Young D. B. Tuberculosis: a problem with persistence. *Nat. Rev. Microbiol.* **2003**, 1, 97-105.

74. Cosma C. L., Sherman D. R., Ramakrishnan L. The secret lives of the pathogenic mycobacteria. *Annu. Rev. Microbiol.* **2003**, 57, 641-676.

75. Nathan C., Shiloh M. U. Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proc. Natl. Acad. Sci. USA* **2000**, 97, 8841-8848.

76. Wayne L. G., Sohaskey C. D. Nonreplicating persistence of *Mycobacterium tuberculosis*. *Annu. Rev. Microbiol.* **2001**, 55, 139-163.

77. Sherman D. R., Voskuil M., Schnappinger D., Liao, R. Harrell, M. I., Schoolnik G. K. Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding Alpha-crystallin. *Proc. Natl. Acad. Sci. USA* **2001**, 98, 7534-7539.

78. Roberts D. M., Liao R. P., Wisedchaisri G., Hol W. G., Sherman D. R. Two sensor kinases contribute to the hypoxic response of *Mycobacterium tuberculosis*. *J. Biol. Chem.* **2004**, 279, 23082-23087.

79. Saini D. K., Malhotra V., Dey D., Pant N., Das T. K., Tyagi J. S. DevR-DevS is a bona fide two-component system of *Mycobacterium tuberculosis* that is hypoxia-responsive in the absence of the DNA-binding domain of DevR. *Microbiology* **2004**, 150, 865-875.

80. Saini D. K., Malhotra V., Tyagi J. S. Cross talk between DevS sensor kinase homologue, Rv2027c, and DevR response regulator of *Mycobacterium tuberculosis*. *FEBS Lett.* **2004**, 565, 75-80.

81. Park H. D., Guinn K. M., Harrell M. I., Liao R., Voskuil M. I., Tompa M., Schoolnik G. K., Sherman, D. R. Rv3133c/dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **2003**, 48, 833-843.

82. Florczyk M. A., McCue L. A., Purkayastha A., Currenti E., Wolin M. J., McDonough K. A. A family of *acr*-coregulated *Mycobacterium tuberculosis* genes shares a common DNA motif and requires Rv3133c (dosR or devR) for expression. *Infect. Immun.* **2003**, 71, 5332-5343.

83. Parish T., Smith D. A., Kendall S., Casali N., Bancroft G. J., Stoker N. G. Deletion of two-component regulatory systems increases the virulence of *Mycobacterium tuberculosis*. *Infect. Immun.* **2003**, 71, 1134-1140.

84. Malhotra V., Sharma D., Ramanathan V. D., Shakila H., Saini D. K., Chakravorty S., Das T. K., Li Q., Silver R. F., Narayanan P. R., Tyagi J. S. Disruption of response regulator gene, *devR*, leads to attenuation in virulence of *Mycobacterium tuberculosis*. *FEMS Microbiol. Lett.* **2004**, 231, 237-245.

Chapter 4

Exploring the Use of α-Helix Mimetics for Probing the Protein–Protein Interactions of DosR in Preventing Its DNA-binding

Introduction and Summary

Many proteins exert their biological activity by interacting with other proteins. Protein-protein interactions play a pivotal role in many key biological processes, which makes the prevention of these interactions a way of modulating the activity of such proteins. Dormancy Survival Regulator (DosR), a protein that is believed to mediate the transition of *M. tuberculosis* into dormancy, recruits a homodimer interface between A and B subunits for subsequent DNA-binding. It can serve as a good model system to investigate the inhibition of protein–protein interactions and any subsequent effect on the bacterium. More specifically, an α-helix/α-helix interaction incorporated in this interface accounts for over 40% percent of the overall interactions. This prompts us to develop α-helix mimetics with a higher affinity to preferentially bind to the α-helix (α10) of the monomer. This will result in disruption of the dimerization of DosR. Interruption of the subsequent DNA-binding may prohibit *M. tuberculosis* from transforming into dormancy, so that current TB therapeutics could come into play.

Here, we take two approaches to address the problem. In our small molecule approach, we started by adopting Hamilton's terphenyl framework. Substitutions at 3, 2′, 2″ positions allow us to explore the DosR interface systematically. Computer modeling suggests 1-benzyl triazole derivatives will be a better scaffold. Unfortunately, our small molecules only showed inhibition at submillimolar range. The other approach is the

peptide approach. We designed and synthesized a series of linear and cyclic peptides with both natural and unnatural amino acids. By preserving the "TVT" motif at correct positions, we either introduce amino acids with a high α-helix propensity or apply additional constrains to the peptides in order to enhance their α-helix percentage. The best peptide inhibitor we obtained so far demonstrated an IC50 of ~50 μM. For peptide mimetics, 16 mer peptide **9** demonstrated an IC50 in the 100 μM range. Bicyclic peptides 8 mer was also shown to have improved the inhibitory activity over linear control peptides.

Results and Discussions

Strategies for α-helix mimetic. In chapter 3, I have reviewed some of the strategies proposed by other researchers for α-helix mimetics. In this project, we are concentrating on exploration of peptide and small molecule inhibitors for DosR. In the DosR interface, T198, V202 and T205 are critical and contribute to over 40% of the total interaction. They reside on the same side of the α10 helix and align very well. We attempted to design some molecules that are able to reproduce this key interaction by presenting TVT in a similar way. We can present TVT by incorporating their side chains onto small molecule templates with the right conformation. By building a library and introducing some variations, we expand our search to a wider range. An alternative means is to embed TVT into some particular peptides which will preferentially adopt an α-helix conformation. The more helical content the peptides have, the more effective the inhibitors are going to be. A series of 18mer peptides incorporating "TVT" were

synthesized to validate the importance of the triad (**Figure 4.1**). Some mutations were made to increase the α-helix propensity (F203A, L207A), some to enhance solubility (M194T), some to balance the helix dipole and some others to improve hydrogen bonding (R197E, K208R). The experimental results support our initial design concept and enable us to pursue our research goal by the synthesis of peptide and non-peptide small molecule inhibitors.

Assay Development and Inhibition Evaluation. The inhibitors were evaluated by using an electrophoretic mobility shift assay, also referred to as gel shift assay. This technique is generally used to study protein-DNA interactions. This process can determine the binding capability of a protein or protein mixtures to a given DNA. The assay is based on the observation that DNA, when in complex with a protein, migrates through a non-denaturing polyacrylamide gel more slowly than free DNA fragments. Because the gel is stained with ethidium bromide, the control lane with DNA alone will show a clear single band, corresponding to an unbound DNA fragment. Upon protein binding, another band will appear in the protein lane, representing the larger protein-DNA complex that has been formed. In the presence of various concentrations of inhibitors, DosR will bind the inhibitors, leaving some DNA in the unbound state. The affinity of the protein to the DNA can be estimated by the ratio of bound to unbound DNA. This assay is very useful for estimating the association and dissociation constant. A typical gel shift assay pattern is attached in **Figure 4.2**. In our assay, we first incubated the DosR monomer with the inhibitor; after that, DNA is mixed with the protein-inhibitor mixture for gel electrophoresis analysis.

Small Molecule α-Helix Mimetics.

Small Molecule Design. Terphenyl has drawn people's attention over the years because of its robustness, UV sensitivity, and its unique conformations.[1-6] The carbon-carbon single bond that connects the two phenyl ring affords one degree of freedom for terphenyl to adopt different conformations. Hamlton's group is the first to utilize *p*-terphenyl as a scaffold to mimic α-helices.[7-11] Based on those successes, we chose terphenyl as our starting point. With terphenyl being the backbone, we only need to present the side chains of TVT correspondingly to prove the design principle. However, computer modeling will be required to verify the design before any further investigation. Even more interestingly, the geometry of the terphenyl offers a great opportunity to present multiple copies of the key sidechains of TVT, thus could potentially enhance the potency of the inhibitors. We proposed the design of terphenyls with two sets of functionalities (**Figure 4.3**). Considering the molecular symmetry, the presence of multiple copies of the side chain could improve the affinity. We designed terphenyl compounds either with rotational symmetry or with plane symmetry. In this case, the sidechains to be presented are 1-hydroxyethyl, isopropyl and 1-hydroxyethyl respectively. For synthetic simplicity, we used an ethyl group to displace the isopropyl group on the middle ring.

Computer Modeling. To verify the low energy conformation of substituted *p*-terphenyl, we built up a model using 3,2′,2″-trimethyl-*p*-terphenyl using FLO96. Initial low-energy conformer search results point out that the tris-methylterphenyl with the lowest energy is in a staggered conformation and it has torsion angles and distances between substituted groups that are very similar to those residues at *i*, *i*+4, and *i*+7 in an

α-helix (**Figure 4.4**). This result matched Halmiton's modeling results fairly well and we decide to employ *p*-terphenyl as a framework for our project. Since the steric hindrance confines the conformation, the question simply becomes how to present side chains. A model compound has been made to mimic the key residues TVT. The corresponding substitutions at three ortho positions are 1-hydroxyethyl, ethyl, 1-hydroxyethyl respectively, projecting on one side of the helix (refer to **2** in **Scheme 4.1**).

Synthesis of Terphenyl Compounds. Now that we have designed our terphenyl molecules, we need to develop some appropriate chemistry to synthesize the terphenyls. Different synthetic strategies were applied for compounds with single set and double sets of sidechain mimetics. **Scheme 4.2** outlined the synthesis of bi-functionalized terphenyl. A convergent synthetic scheme was developed employing sequential Suzuki coupling reactions. The individual monomer 3-acetophenyl boronic acid and 2-acetophenyl boronic acid are commercially available. These monomers were sequentially assembled on the middle ring by reacting with 1,4-dibromobenzene. Meanwhile, these precursors allow additional substituents to be added to the 3 and 2″-position. On the other hand, compounds with 3, 3″ substitutions are also synthesized (**2a**, **2c**, **2f**, **2h** in **Figure 4.5**). The advantage of these compounds from the synthesis standpoint is they can be synthesized in only one step of reaction. In an attempt to prepare hexa-substituted-terphenyl (**Scheme 4.3**), 2,5-dihydroxy-dimethyl terephthalate was chosen as the starting material and we tried the reactions under a variety of standard coupling conditions. However, the reaction was not successful on account of the tremendous steric hindrance as well as the low eletrophilicity of the aromatics. Since it has been widely recognized that Suzuki coupling under sterically demanding conditions or with electro withdrawing

groups dramatically reduces the reactivity,[12-16] we modified our design accordingly to C in **Figure 4.3**. The 2,5-diethyl-1,4-bisboronic acid was ready prepared from 1,4-dibromo-2,5-diethylbenzene.[17] The bis(neopentyl glycolato) boronic ester was hydrolyzed to give the corresponding boronic acid. Subsequent Suzuki coupling with triflates under the existence of $Pd(PPh_3)_4$ produced the 3,2″-bisaceto-*p*-terphenyl. The final product was obtained after reduction with either $NaBH_4$ or $LiAlH_4$. A small library has been synthesized using the method described herein (**Figure 4.5**).

Screening of Terphenyl Compounds. A small library of terphenyl compounds were synthesized and screened in the gel shift binding assay. Since the terphenyl scaffold is highly hydrophobic in nature, up to 40% of DMSO was used together with the buffer solution to enhance the solubility. Compound **2a** and **2b** demonstrated inhibition at low micromolar range (**Table 4.1**) but was later attributed to protein precipitation. To our surprise, compound **2c** with 1-hydroxyethyl at the desired position showed a very poor IC50; whereas compound **2d**, which has the side chain of threonine at 3″ exhibited an IC50 of around 250 μM. As we anticipated, terphenyl inhibitor **2i** with an ethyl group on the middle ring, mimicing the valine side chain, displayed an affinity gain of at least 10 fold compared with **2c**. However, SDS-PAGE results raised the concern of possible protein precipitation across this series (**B** in **Figure 4.6**). As the inhibitor concentration increased, less amount of $DosR^C$ was observed, which made our IC50 not very conclusive. Some modifications need to be made to enhance the solubility as well as the binding affinity.

Virtual Screening of Terphenyl Inhibitors. In order to discover the most promising fragments for the binding site and further optimize the existing inhibitors, we

conducted virtual screening with FLO96,[18,19] a useful tool for docking and molecular modeling. We built a 10×5×10 library (**Figure 4.7**). The top 200 hits were visually inspected to verify their legitimacy. The results differ considerably from our perception —— the 1-phenylethanol portion of the terphenyl bound snugly to the binding pocket, while the rest of the molecule either bound to a different pocket or totally pointed away (**Figure 4.8**). The screening results suggest that terphenyl could not deliver the second residue to the correct location. In another word, the terphenyl scaffold is not long enough to reach the hydrophobic pocket regardless of the substitutions. Terphenyl's high hydrophobicity narrowed its application, particularly in this case, where DosR has some hydrophilic character in the binding pocket.

Iterative Inhibitor Design and Synthesis. We scrutinized the crystal structure of $DosR^C$ and it became quite clear that if we retain the 1-phenylethanol moiety, a one-carbon linker is needed to connect the second ring in order that it can fit in the right pocket. A two-carbon tether will be ideal for the third moiety to be in position. Considering this new information, we constructed a new scaffold in FLO96 (compound **3** in **Scheme 4.1**, also **Figure 4.9**). Besides the right positioning, this new molecule has one more hydrogen bond donor on the second ring, interacting with Thr205. A salt bridge at the far end of the third ring, interacts with Glu206, making additional contribution to the total stabilization energy.

Due to the restriction of chemical availability, we made an analogous compound **3c** instead. This compound embraces the essential features of **3** and we use it as a model compound for inhibition study. We employed "click chemistry" into our synthesis. This method allows connections of olefins, electrophiles, heteroatoms in a fast and modular

way for efficient synthesis.[20-22] 3-bromomethyl methyl benzoate reacted with sodium azide in DMF for 2 h to give the corresponding azide **3a**. Subsequent reaction with 3-phenylpropyne with sodium L-ascorbate produced compound **3b**. Sodium boronhydride reduction provided the target compound **3c** (**Scheme 4.4**). Unfortunately, compound **3c** did not demonstrate inhibition at 1 mM. The unsuccessful exploration in the small molecule realm prompted us to shift our effort for searching peptide inihibitors for DosR.

Peptide α-Helix Mimetics

Efforts to develop peptide therapeutics have been growing in the past few years.[23-25] Biological diversity alongside high specificity, affinity and molecular recognition are all important peptide virtues. Another reason why peptide drugs are catching so much attention is because its fewer toxicology issues, which has been the graveyard of so many small molecules. We can rationally and systematically optimize peptides. On the other hand, people are developing new technology to address peptide shortcomings.[26, 27]

In the DosR project, we are in search of specific peptides that are able to preferentially adopt an α- helix conformation. The high helix propensity and better solubility will give the peptide inhibitors better chances of success as compared to their small molecule counterparts. It has been widely accepted that short peptides tend to form random coils instead of folding into regular secondary structures in solution. Therefore, our question becomes how to increase the helical content of peptide in the aqueous solution. Presently, several strategies have been reported: 1) α-helix can be stabilized either by pushing the random coil/α-helix equilibrium towards helix formation,[28] or 2) stabilizing the dipole of the helix,[29] or 3) incorporating amino acids with higher α-helix

propensity, or 4) including nucleating templates,[30] or 5) applying structural constraint on the peptide.[31-34] In this study we search peptide inhibitors of DosR in two way —cyclic peptides and linear peptides. Cyclic peptides were obtained by side chains cyclization with lactam bridges. For linear peptides, we incorporated both natural and unnatural amino acids and amino acids with high helix propensity to enhance helicity. Although it has been realized that introduction of pairs of D-amino acids tend to disrupt proximal secondary structures,[35] they can be used to retain the intrinsic hydrophobic content of the peptide, particularly in our case, extra constraints will be applied, which would possibly lock the peptide's conformation.

Peptide Design. Among the many ways to constrain peptide backbone via cyclization, we used the side chain to side chain cyclization to force peptides into bioactive conformations. In a typical α-helix, the minimal repeating unit is defined by the hydrogen bonding between carbonyl oxygen from one residue (position *i*) and amide NH from another residue (position *i*+3/*i*+4). The simplest method is to enhance the helical content by forming intromolecular linkage between residues *i*, *i*+3 or residues *i*, *i*+4. We choose to cyclize side chains of residues *i*, *i*+3 with the shortest linkage. In proteins, all amino acids are L-amino acids. If we can change the amino acid at *i*+3 position to a D-amino acid, then the side chain would point towards each other. As a result, the linker between *i*, *i*+3 can be shortened compared with L-amino acids. With truncation, we could find some short peptides with a high percentage of α-helical content.

Computer Modeling. We used FLO96 to run a conformer search and find out the optimal linkage between the two residues. By varying the length of the linker

systematically, we identified that a seven-carbon chain (or equivalent chains with heteroatoms) is required between the residues *i* and *i*+4 to stabilize the α-helix. We also found out the right linker for cyclization at *i* and *i*+3 positions, but the peptide is either deflected from helix or totally distorted. We envisioned a D-amino acid at *i*+3 position may lead to a much shorter linker. The results were encouraging and the helix is well stabilized by a four-carbon chain (**Figure 4.10**). For *i*, *i*+3 cyclization, candidates for the chemistry are rather limited. Disulfide bond is a good choice because it widely exists in the biological system to stabilize tertiary or quaternary structures. The reaction condition is very mild and the cyclization can be finished off resin.[36] Cysteine provides the side chain with the correct side chain length. Another possible choice is the amide bond. The chemistry is very clean and should be completed on resin with a high yield. Aspartic acid and L-α,β -diaminopropionic acid (Dap-OH) will afford COOH and NH_2 respectively. A third option is triazole linkage. Although it seems to surmount the optimal number of carbons, its compact structure compensates for it. We submitted our selections to FLO96, the modeling results substantiated our predictions —— all the above linkers favor helix formation.

Peptide Synthetic Strategy. We actively pursue the amide bond and disulfide bond formation because of material availability. With our previous design principles well established, we are now trying to truncate the peptide furthermore. At the same time, we introduced constrains by incorporating the D-Asp and L-Dap pair at *i* and *i*+3 positions of an α- helix. This leads to intramolecular amide bond formation with little disturbance on helix formation. An 8mer peptide is the theoretical minimum to present TVT in the right way. Peptides **4-7** (**Figure 4.11**) were chosen as models to investigate

the helix-stabilizing potential over a span of eight residues. Peptide **4** is a direct truncated version of native DosR α10 helix and peptide **5** incorporates the linker amino acids (D-Asp and L-Dap) but remains linear. Peptide **6** has a single lactam ring and peptide **7** is doubly cyclized. A tyrosine was added to the peptide sequence for characterization purposes.

All cyclic peptides are synthesized on solid support using Fmoc-chemistry with orthogonal protecting groups [37] (**Scheme 4.5**). The N-terminals of the peptides are capped with acetyl group, because it has been reported to have a profound effect for α-helix stabilization.[38] For peptides with single cyclization, we used allyl protected Asp and alloc protected Dap, because they are orthogonal-protected building blocks, thereby permitting the synthesis of cyclic peptide without affecting the main chain propagation. The combination of allyl and alloc is particularly advantageous since both side chains can be simultaneously unmasked in a single step with $Pd(PPh_3)_4$ as a catalyst. For peptides with double cyclization, we employed another orthogonal protection pair——2-phenylisopropyl and 4-methyltrityl. They are super acid sensitive protecting groups and can be easily removed under mild acidic conditions (1% TFA in CH_2Cl_2) without interfering with other protected side chains. In the case of the disulfide bond linker formation, we used trityl protected L- and D-Cysteine, taking advantage of their acid-labile nature. The product is yet to be purified by HPLC.

In our effort to obtain linear peptide inhibitors with high helicity and potency, we recruited some proteinogenic as well as nonproteinogenic amino acids which induce the helix formation, for example, 2-Aminoisobutyric acid (Aib),[39-41] Glutamate, Lysine,[42] etc (**Figure 4.12**). We also mutated some hydrophobic residues with charged residues to

facilitate proper peptide solubilization and helical dipole stabilization. Because of the intrinsic steric hindrance of Aib, the coupling reaction was repeated over an extended period of time.

IC50s. Gel-shift assay results denoted that all the peptides showed a large increase in inhibition upon cyclization. The eight amino acids control peptide **4** showed no inhibition at 1mM; monocyclic peptide **6** showed an IC50 of ~400 μM and bicyclic peptide **7** gained some more affinity and had an IC50 of ~200 μM, which accords well with our initial design. Longer peptides containing "TVT" were scrutinized to improve solubility, a bottleneck for the 8mers. 16mer peptide **9** and **10** were derived from 18mer peptide **2**. Peptide **9** behaved very similarly to peptide **2** and exhibited an IC50 of ~100 μM. Peptide **10** which contains multi Aib showed, to our surprise, no significant improvement in inhibition compared with peptide **9**.

Helical Content Determination.

Circular Dichroism. The α-helicity of the peptides was determined by measuring their characteristic ellipticity at 208 nm and 222 nm in circular dichroism (CD) spectra. A slight shift in the double minima is expected for shorter peptides.[43] The most valuable information that can be extracted from CD is that about the molecule's secondary structures. Following the formulism developed by Chen et al [44] and optimized by Luo and Baldwin,[45] quantitative analysis of mean helix content (f_H) can be achieved.

$$f_H = ([\theta]_{obs222} - [\theta]_C)/([\theta]_{\infty 222} - [\theta]_C) \quad (1)$$

Among which, $[\theta]$ is the mean residue weight ellipticity, $[\theta]_C$ and $[\theta]_{\infty 222}$ are the random coil and infinite α-helix molar ellipticities. $[\theta]_C$ is a linear function of

temperature T (in Celsius) $[\theta]_C = 2220 - 53T$ (2)

$$[\theta]_{\infty 222} = (-44000 + 250T)(1-k/N) \quad (3)$$

Where N is the number of peptide units, and k is a semi-empirical factor correlated to the chain length. While equation 1 has been widely acknowledged, there has been some disagreement on the implementation of proper k values, which covers a range from 2.4 to 4.5.[46] This inconsistency is usually caused by different temperature, additive and modification of the peptide. Baldwin suggested that k equals 3.0 for peptides that are carboxyamidated and 4.0 for those without capping. Given this situation, it may be more useful to calculate relative helicities under very similar experimental conditions rather than absolute mean helix content, especially considering the wavelength shift in the double minima. Here, we calculated the relative helicity percentages for peptide 8mers in **Table 4.2**, and peptide 16mer and 18mer in **Table 4.3**, assuming peptide **4** and **10** to be 100% helical respectively in 50% TFE. Mean residue ellipticities were converted from ellipticity readings using equation 4.

$$[\theta] = \theta_{obs}/10 \times C \times N \times l \quad (4)$$

where θ is the ellipticity in milidegrees, C is the molar concentration of the peptide, N is the number of residues in the peptide, and l is the path length of the cuvette in centimeter. 2,2,2-trifluoroethanol (TFE) is believed to favor α-helix stabilization in solution.[45] It was found out later that the optimum condition is 50% TFE in water (v/v).

Our systematic study of the CD spectra of peptides **2**, **4**, **5**, **6**, **7**, **9** and **10** (see **Figure 4.11 & 4.12** for sequences) are shown in **Figure 4.13-4.16**. The18mer control peptide **2** showed about 25% helical content in H_2O. However, it turned out that the helical content escalated significantly in 50% TFE/H_2O (**Figure 4.13**). Under the same

conditions, 8mer control peptide **4** demonstrated a similar accretion in 50% TFE. Monocyclic and bicyclic peptides **6** and **7**, however, did not show a double-minimum in either solution (**Figure 4.14**). Nevertheless, there were no apparent changes detected in different solvent systems, indicative of the secondary structure of the peptides may have been locked. This result agrees well with our initial speculation of the constraint peptides. In another word, we have stabilized our peptides' secondary structure. It has been reported recently in the literature that D-amino acids tend to cancel out the CD signal of L-amino acids.[47,48] This is in good agreement with our findings and it very well explains the disappearance of the double-minimum.

For the linear peptide, we successfully achieved helicity enhancement from 18mer to 16mer. Shorter peptides tend to adopt random coil. Both 16mer **9** and **10** demonstrated higher helical content in both solutions (**Figure 4.16**).

2-Dimensional (2D) NMR Spectroscopy. In order to verify the helical content of our synthesized peptides under experimental conditions, we did some two-dimensional NMR studies. Traditionally, people use 2D ^{1}H NMR to analyze the connectivity of H-N and N-N, H-H between or within residues.[49,50] That usually could not provide very much quantitive information about secondary structure. Here, we would exploit the unique chemical shift deviation patterns of the secondary structure to quantitate α-helical content in given peptides. The chemical shifts of the synthetic peptides were compared with random coil ^{1}H chemical shifts for the 20 common amino acids residues.[50] 2D TOCSY spectra at 500 MHz were used to identify each individual residues in the peptide sequences (**Figure 4.17**). Owing to the small molecular size, 2D-ROESY was utilized instead of NOESY. 2D TOCSY had to combine with 2D ROESY to allow assignment of

specific residues due to their repeated occurrence in the sequence (namely, Thr and Ala). Almost all residues in the cyclic peptide **6** displayed an upfield shift (0.05-0.51 ppm) relative to random coil (**Figure 4.18**), characteristic of an α-helix conformation.[51-53] Because there are no random coil chemical shift data available for D-amino acids, the chemical shift deviation (CSD) of Asp6 was not calculated. Peptide **4**, which is a 8mer truncation of the α10 native sequence of DosR, did not demonstrate consistent negative CSD; for residues with negative CSD, the deviation was very small (0.01-0.06 ppm, **Figure 4.19**). The consistently negative chemical shift deviation indicates there is more α-helical content in cyclic peptide **6** than in linear peptide **4**. The 2D ^{1}H NMR of bicyclic peptide **7** is in progress. So far, out of the 6 residues that are possible for assignment, 4 residues have demonstrated consistently negative chemical shift (data not shown), characteristic of an α-helix.

Conclusions

In this project, we attempted to synthesize efficacious α-helix mimetics targeting α10 helix of the DosR-DosR interface. This goal was accessed in both the small molecule approach and the peptide approach. In our small molecule efforts, we started with the terphenyl scaffold. We did obtain some inhibitor at ~40 μM range by presentating the TVT side chains on the 3, 2′, 2″ positions. However, due to the potential solubility and other issues brought about by terphenyl's hydrophobic nature, we modified our scaffold to allow the second key functionality to be positioned to the right pocket. Unfortunately, the compound maintaining the most part of the designed molecule did not show inhibition. In our peptide approach, cyclization at *i* and *i*+3 positions with an L-Dap and D-Asp stabilized α-helix conformation of the peptide thus improved the performance of the peptide inhibitors. Even though some other cyclization chemistry is also available, we concentrate on the lactam ring. The increase in cyclic peptide helicity was supported by 2D ^{1}H NMR. In order to further favor α-helix formation and enhance peptide solubility, we extended our search to longer linear peptides, which include unnatural amino acid Aib and α-helix inducing natural amino acids. Following the lead 18mer peptide **2**, a potent peptide discovered in an earlier study, we made 16mer peptide **9** and **10**. By stabilizing the dipole moment at both N and C terminus, we successfully improved the helicity according to the CD spectra. Cyclization of the 16mer peptides using the *i*, *i*+3 cyclization strategy developed in this study might have a better chance of getting a micromolar inhibitors. More important, the results from this study open up some more opportunity to systematically explore the DosR-DosR interface.

Experimental Section

General. All chemical reagents for small molecule synthesis were purchased from Aldrich, except dimethyl chloroterephthalate was purchased from Matrix Scientific, 1,4-dibromo-2,5-diethylbenzene from TCI and [1,1'-Bis(diphenylphosphino)ferrocene] from Frontier Scientific. All amino acids and peptide building blocks were purchased from Novabiochem, SynPep Corporation, or NeoMPS Inc., except D-Met(Trt) was purchased from Matrix Scientific. Poly-Prep chromatography columns for solid phase synthesis were purchased from Bio-Rad. All solvents used were obtained from Fisher Scientific. All reagents and common solvents were of reagent grade or HPLC grade and were used for synthesis and separation without further purification unless otherwise noted. Microwave syntheses were conducted on a CEM discover microwave synthesis system. Thin layer chromatography was performed on J. T. Baker TLC Plates (silica gel on plates, 250 μm layer thickness with a fluorescent indicator). Silica gel used in column chromatography was 70-230 mesh. HPLC purification was performed using a preparative Hewlett-Packard 1100 series HPLC system with a variable wavelength detector. The Columns used is a ZORBEX 300 SB-AQ reverse-phase column from Agilent (21.2 mm × 150 mm, 5 μm). ^{1}H-NMR spectra were recorded at 300 MHz on a Bruker AV300. 2D ^{1}H NMR and 1D ^{13}C NMR spectra were recorded at 500 MHz on a Bruker AV500. Mass spectra were obtained using an Agilent 1100 series LCD/MS mass spectrometer equipped with an ion trap.

Synthesis of Terphenyl Derivatives. *Synthesis of 3, 3″-di-(1-hydroxyethyl)-p-terphenyl (**2c**).* 125.0 mg (0.5 mmol) of 1,4-dibromobenzene, 500 mg (3.0 mmol, 6 eq.)

of 3-acetylphenyl boronic acid and 80 mg (70 μmol, 0.015 eq.) of tetrakis(triphenylphosphine)palladium(0) were dissolved in 4.5 mL of 9:1 DME/EtOH (DME = 1,1-dimethoxyethane). To this yellow solution 0.5 mL of 2M (1.0 mmol, 2.0 eq.) Na_2CO_3 aqueous solution was added. The resulting mixture was heated at 85°C for 5 min under microwave radiation. After being concentrated in vacuo, the residue mixture was taken up with water and extracted with EtOAc. The combined organic fractions were dried over $MgSO_4$ and evaporated. The purified (Flash chromatography EtOAc/Hexane 1/19, Rf = 0.7) and dried di-acetoterphenyl precursor (**2a**) was reduced with 380 mg of $NaBH_4$ (5.0 mmol, 10 eq.) in anhydrous CH_2Cl_2 at room temperature for 2 hr (**Scheme 4.2**). HPLC yielded 129 mg of product (0.405 mmol, 81%) as clear oil after lyophilization. MS (**2c**): 318.2. ^{1}H NMR of **2c** ($CDCl_3$): δ 1.58 (6H, d), 1.77 (2H, s), 5.02 (2H, q), 7.43 (2H, d), 7.48 (2H, t), 7.60 (2H, d), 7.66 (6H, m). MS (**2a**): 314.1. ^{1}H NMR of **2a** ($CDCl_3$): δ 2.68 (6H, s), 7.43 (2H, m), 7.51 (4H, m), 7.56 (2H, m), 7.78 (2H, m), 7.98 (2H, m).

*Synthesis of 3, 2″-di-(1-hydroxyethyl)-p-terphenyl (**2d**).* 288.0 mg (1.2 mmol) of 1,4-dibromobenzene, 209 mg (1.2 mmol, 1 eq.) of 2-acetylphenyl boronic acid and 35 mg (30 μmol, 0.025 eq.) of tetrakis(triphenylphosphine)palladium(0) were dissolved in 4.5 mL of 9:1 DME/EtOH. To this yellow solution 1.0 mL of 2M (2.0 mmol, 2.0 eq.) Na_2CO_3 aqueous solution was added. The resulting mixture was heated at 85°C for 5 min under microwave radiation. After subsequent evaporation and extraction, the produced biphenyl bromide was mixed with 209 mg (1.2 mmol, 1 eq.) of 3-acetylphenyl boronic acid and 35 mg (30 μmol, 0.025 eq.) of tetrakis(triphenylphosphine)palladium(0),

4.5 mL of 9:1 DME/EtOH and 1.0 mL of Na_2CO_3 aqueous solution. The resulting mixture was heated again under microwave irradiation for 5 min. After being concentrated in vacuo, the residue mixture was taken up with water and extracted with EtOAc. The combined organic fractions were dried over $MgSO_4$ and evaporated. The crude di-acetoterphenyl precursor (**2b**) was first purified by flash chromatography (EtOAc/Hexane 1/4, Rf = 0.4) and then reduced with 380 mg of $NaBH_4$ (5.0 mmol, 5 eq.) in anhydrous CH_2Cl_2 at room temperature for 2 hr. Clear oil was obtained after HPLC purification. MS (**2d**): 318.2. ^{1}H NMR of **2d** ($CDCl_3$): δ 1.47 (3H, d), 1.59 (3H, d), 3.52 (2H, s), 5.06 (2H, dq), 7.33-7.50 (8H, m), 7.58-7.79 (4H, m). MS (**2b**): 314.1. ^{1}H NMR of **2b** ($CDCl_3$): δ 2.64 (6H, s), 7.33-7.43 (2H, m), 7.56-7.60 (6H, m), 7.68 (1H, m), 7.72--8.00 (3H, m).

*Synthesis of 2′,5′-diethyl-3,3″-di-(2-hydroxyisopropyl)-p-terphenyl (**2e**).* Precursor **2b** (78.5 mg, 0.25 mmol) and methylmagnesium bromide (3M solution in Et_2O, 0.8 mL, 10 eq.) were dissolved in 10 mL of anhydrous THF at -78°C with N_2. The reaction was allowed for 1 hr, followed by hydrolysis with 2M HCl (aq.) for 30 min. The resulting solution was evaporated and the residual extracted with Et_2O. After evaporation, the crude product was purified by HPLC. MS: 346.2. ^{1}H NMR ($CDCl_3$): δ 1.56 (12H, s), 7.15-7.25 (5H, m), 7.38-7.41 (3H, m), 7.50-7.54 (4H, m).

*2′,5′-diethyl-3,3″-di-(1-hydroxyethyl)-p-terphenyl (**2f**).* MS: 370.2. ^{1}H NMR ($CDCl_3$): δ 1.11 (6H, t), 2.56-2.66 (10H, m), 7.17 (2H, s), 7.52-7.62 (4H, m), 7.96-7.99 (4H, m).

*2',5'-diethyl-3,2"-di-(1-hydroxyethyl)-p-terphenyl (**2g**).* MS: 370.2. ^{1}H NMR ($CDCl_3$): δ 1.10 (6H, dt), 2.07 (2H, s), 2.44-2.62 (4H, m), 2.63-2.68 (4H, m), 7.06 (1H, s), 7.16 (1H, s), 7.33 (1H, d), 7.35-7.54 (4H, m), 7.73 (1H, d), 7.97 (2H, m).

*2',5'-diethyl-3,3"-di-(1-hydroxyethyl)-p-terphenyl (**2h**).* 2h was synthesized following a similar method as described above for **2c** and **2d**. MS: 374.2. ^{1}H NMR ($CDCl_3$): δ 1.14 (6H, t), 1.58 (6H, d), 2.63 (4H, q), 4.99 (2H, q), 7.18 (2H, s), 7.28-7.34 (4H, m), 7.41-7.48 (4H, m).

*2',5'-diethyl-3,2"-di-(1-hydroxyethyl)-p-terphenyl (**2i**).* MS: 374.2. ^{1}H NMR ($CDCl_3$): δ 1.12 (6H, dt), 1.40 (3H, d), 1.58 (3H, d), 1.76 (4H, s), 2.33-2.45 (2H, m), 2.46-2.64 (2H, m), 4.71-4.87 (1H, dq), 5.00 (1H, q), 7.16-7.21 (2H, m), 7.28-7.73 (6H, m).

*Synthesis of 1,4-diethylbenzene-2,5-bis(boronic acid) (**2m**).* 100 mg (0.34 mmol) of 1,4-dibromo-2,5-diethylbenzene, 134 mg (0.82 mmol, 2.4 eq.) of bis(neopentylglacolato) diboron, 7.3 mg (0.01 mmol, 0.03 eq.) $PdCl_2$dppf, and 74.5 mg (1.0 mmol, 3 eq.) potassium acetate were mixed in anhydrous DMSO and heated at 80°C for 4 hr. The reaction progress was checked by TLC. 2M HCl aq was added to the reaction mixture for hydrolysis (45 min). The reaction solution was cooled down to room temperature and poured into ice-water. After extraction with EtOAc, the combined organic solution were dried over $MgSO_4$ and evaporated in vacuo. The residue was purified by HPLC and white powered was obtained at over 85% yield. MS: 222.1. ^{1}H NMR (D_2O): δ 1.11 (6H, t), 2.67 (4H, q), 7.17 (2H, s), 7.93 (4H, s).

Attempted Synthesis of Symmetric Terphenyl A Derivative (A). 254 mg (1.0 mmol) of diethyl 2,5-dihydroxyterephthalatewas dissolved in anhydrous DMF. The resulting solution was added dropwise to a suspension of K_2CO_3 in THF (210.0 mg, 1.5 mmol, 1.5 eq.). After stirring at room temperature for 2 hr, a solution of benzyl bromide (0.18 mL, 1.5 mmol, 1.5eq.) in THF was added to the mixture dropwise. Potassium Iodide (33.0 mg, 0.2 mmol, 0.2 eq.) was also added to accelerate the reaction. The reaction was allowed for 3 hr and was quenched with saturated sodium bicarbonate aqueous solution. The solution was extracted with EtOAc twice and the combined organic solution was evaporated and dried over $MgSO_4$. The monobenzylated product **2k** was purified by flash chromatography (EtOAc/Hexane 15/85). Rf = 0.7. MS: 316.1. ^{1}H NMR ($CDCl_3$): δ 1.34 (3H, t), 1.43 (3H, t), 4.35 (4H, dq), 5.11 (2H, s), 7.26-7.50 (7H, m). To 127 mg (0.5 mmol) of **2k** is dissolved in redistilled pyridine was added triflic anhydride (1.66 mL, 10 mmol, 20 eq.). The reaction was allowed at -78°C under nitrogen for 8 hr. The reaction was quenched with $NaHCO_3$ aqueous solution and extracted with EtOAc. The combined organic layer was evaporated and dried over MgSO4. The crude product **2l** was purified by flash chromatography (EtOAc/Hexane 20/80). Rf = 0.3. MS: 448.1. ^{1}H NMR ($CDCl_3$): δ 1.34 (3H, t), 1.43 (3H, t), 4.38 (2H, q), 4.46 (2H, q), 5.23 (2H, s), 7.35-7.50 (4H, m), 7.70-7.71 (2H, m).

Synthesis of Benzyltriazole Derivatives. A flask was charged with 138 mg of 3-bromomethyl methyl benzoate (**3**, 1.0 mmol), 325 mg of sodium cyanide (5.0 mmol, 5 eq.) and anhydrous DMF. The reaction was allowed at room temperature for 4 hr. The reaction progress was monitored by TLC. The reaction was quantitative. 38.2 mg **3a** (0.2 mmol), 26.1 mg (28.2 μL, 0.2 mmol, 1 eq.) 3-phenylpropyne, 8.0 mg sodim L-

ascorbate (0.02 mmol, 0.1 eq) and 5.0 mg (0.04 mmol, 0.2 eq.) $CuSO_4 \cdot 5H_2O$ were mixed in anhydrous DMF. The resulting mixture was heated at 90°C for 10 min. The mixture was then cooled to room temperature and dissolved in EtOAc. The organic layer was washed with water, 2M HCl, $NaHCO_3$, and brine sequentially. After evaporation, the crude **3b** was purified by flash chromatography (EtOAc/Hexane 15/85). Rf = 0.65. All the obtained **3b** was redissolved in anhydrous THF together with 76 mg (2.0 mmol, 10 eq.) $NaBH_4$, and the reaction was allowed at room temperature for 2 hr, followed by 1 hr of hydrolysis under 1M HCl (**Scheme 4.4**). MS (**3c**): 293.2. ^{1}H NMR ($CDCl_3$, **3c**): δ 2.87 (4H, s), 4.75 (2H, s), 4.99 (2H, s), 6.98-7.01 (4H, m), 7.08-7.12 (3H, m), 7.21 (2H, t), 7.30 (1H, s).

Gel-shift Assay. The DNA Gel-shift assay of DosR was run with a 43mer DNA probe on a 15% TBE gel, which was precast for polyacrylamide electrophoresis. 1.0 μL of 50 mM Tris buffer, 2.0 μL of 5.0 μM $DosR^C$, and 10.0 μL of inhibitor solution of various concentrations were mixed together and incubated at room temperature for 30 min. After that, 2.0 μL of 2.5 μM DNA, and 5.0 μL of loading dye (5× stock: 200 μL 10×TBE, 400 μL of glycerol, 400 μL of 6× DNA loading buffer) were added to the solution sequentially. The final concentrations of most inhibitors are 20 μM, 40 μM, 100 μM, 200 μM, and 1000 μM respectively. The dye appeared blue, indicating pH 8. Upon the addition each component, the solution is suspended ~15 times. After all the components are mixed together, the eppendorf tubes are centrifuged at 14 krpm for 5 min. 5.0 μL of the mixture was loaded onto a 15% TBE polyacrylamide gel and

electrophorsed at 200 V for 45 min. The gel was stained by ethydium bromide and visualized with a 315 nm transilluminator.

Computer Modeling. *Virtual Screening.* Docking was carried out by using the CombiDOCK [53] algorism for the virtual screening of the combinatorial library (**Figure 4.7**). CombiDOCK is a simple variation of DOCK, preserving most features of DOCK and its adaptation for the combinatorial library makes it especially useful for fragment docking at all sites. The combinations of fragments are made and only the best combinations are checked for internal clashes and saved if no clashes are found. The strength of the program is that the combinatory process can be achieved by simple addition of the score at each fragment. Conformational search is conducted by energy minimization procedure of FLO96. Each compound was subjected to 1,000 cycles of conformer search and energy minimizations to identify conformations most likely to bind to the enzymatic pocket defined (by Thr198, Ala201, Val202, Thr205). Each cycle involved 400 rapid Monte Carlo search steps followed by energy minimization of the best conformer from the set of 400. The program returned the 2000 lowest energy conformations and they were divided into 20 subgroups. Twenty five conformers with the highest score in each subgroup were visual inspected.

Iterative Design of New Inhibitors. Compound **3c** was constructed under FLO96. The first benzene ring was conserved because of its tight interaction with Gly164 and Arg197. A pyrrolidine moiety was added to form a hydrogen bond with Thr205 with a one-carbon linker. A six-member ring was also incorporated into the new molecule due to its hydrophobic interaction with Val202 and Thr205, which is two carbons away from the second ring. Finally, another pyrrolidine ring was added to the end of the molecule to

form a salt-bridge with Glu206 (**Figure 4.9**). Energy minimization of **3c** towards the binding pocket was conducted by Dockmin.

Conformer Search for Peptido α-helix Mimetics. Conformer search is by energy minimization procedure of FLO96. The bicyclic peptides were subjected to 1,000 cycles of conformer search and energy minimizations to identify conformations most likely to bind to stabilize an α-helix. Each cycle involved 400 rapid Monte Carlo search steps followed by energy minimization of the best conformer from the set of 400. The program returned the 25 lowest energy conformations and conformer with the lowest energy matched the poly-alanine helix perfectly (**Figure 4.10**).

Peptide Synthesis and Purification. Linear and cyclic peptide analogues were synthesis through standard Fmoc-chemistry on the solid support. Assembly of the peptide chain was carried out on rink amid MBHA resin with PyBOP activation. All synthetic peptides were acetyl (Ac) blocked at the N-terminus and amidated at the C-terminus. Cleavage of the peptides was performed using a TFA in the presence of TIS and H_2O (94/3/3) with concomitant removal of the protection groups. The TFA solution was evaporated in vacuo and washed with cold ether to remove the organic residual. The resulting precipitate was washed with cold ether for at least two times. The peptide was then redissolved in MeOH/H_2O mixture for purification. After evaporating the solvent, the pure peptides were lyophilized for further applications. Cyclization of the lactam ring of the linear peptides occurred on the rink amide MBHA resin after the chain propagation was completed. The acid-labile protection groups Mtt and 2-PhiPr were removed with 1% TFA in CH_2Cl_2 (2 min each, four times). Subsequent on-resin cylclization was performed with PyBOP/HOBt chemistry. The orthogonal protection groups Allyl/Alloc

were removed by $Pd(PPh_3)_4$. Peptides were purified by C18 reverse phase-HPLC. All peptides were confirmed by $(M+H)^+$, or $(M/2+H)^+$, $(M/3+H)^+$ and $(M/4+H)^+$. Mass spectrum of peptide 10 is shown in **Figure 4.20** as a typical example.

Circular Dichroism. Circular dichroism (CD) experiments were performed on a Jasco Model J-720 circular dichroism spectrophotometer (Jasco Inc.). Spectra were recorded in a 0.1 cm cuvette (Hellma Cells Inc.) between 270 nm and 190 nm at 100 nm/min with a bandwidth of 2.0 nm, response time of 1 s, and resolution step width of 0.2 nm. Each spectrum is averaged on 12 scans with smoothing to remove noise using Spectra Analysis (version 1.52.01, Jasco Inc.). Peptide samples were dissolved in deionized water at ~1 mM. Final sample was prepared by serial dilution in a range of 15 μM to 400 μM. In cases where TFE was utilized as an additive, peptide samples were first dissolved in H_2O and then diluted with 50% TFE aqueous solution. Peptide samples were measured at various concentrations with independent runs. All spectra were recorded at 20°C. The temperature control was achieved by using a Neslab circulating water bath and a PFD-425S (Jasco Inc.) thermostated cell holder.

2-Dimensional NMR. 2D TOCSY spectra were recorded at 298 K on Bruker AV-500 and DRX-499 spectrometers equipped with ^1H two gradient probes. For TOCSY, a mixing time of 60 ms was used and for ROESY 300 ms was employed. 1024 complex point along the F2 dimension and 256 complex points along the F1 dimension were used for recording TOCSY and 1024 complex points and 512 complex points were used along F2 and F1 axes respectively for the ROESY spectra. (When recording both the spectra, the WATERGATE solvent suppression scheme was used to eliminate the dominant HOD resonance.[55-59]) 2D ^1H NMR spectra were processed using XWINNMR (Bruker). A

shifted sine bell window function was used along both dimensions and the data were zero filled to obtain 1024×1024 data matrix followed by Fourier transformation (FT) and high-order polynomial base line correction. Chemical shifts were referred to DSS, an internal standard at 0.00 ppm. Processed data was further analyzed using NMR assignment and intergration software Sparky [60] with the sequencial assignment method developed by Wuthrich.[49]

Peptide samples for 2D ^{1}H NMR analysis were prepared by dissolving 1.5 to 2.0 mg of peptide in 500 μL solution mixture, which is composed of 450 μL of 50 mM phosphate buffer, 35 μL of D_2O and 15 μL of 3 mM DSS (DSS = 2,2-dimethyl-2-silapentane- 5-sulfonate sodium salt) in D_2O. The sample was then transferred into a 528-PP (Wilmad LabGlass) NMR tube for data collection. The pH of all samples is around 7.

	193	198	202	205
DosR α10		GMERRTQAAVFATELKR		
Peptide1		YGTERETQAAVAATEARR		
Peptide2		YG[illegible]ERETQAAV[illegible]ATE[illegible]RR		

Figure 4.1. The sequences of the peptide 18mers. The first line is the native sequence truncated from α10 helix. Mutations in peptide **1** and **2** are made to either increase the α-helix propensity or to stabilize the hydrogen bonding, or to enhance the solubility. TVT are colored in red, mutations are in blue and cyan respectively.

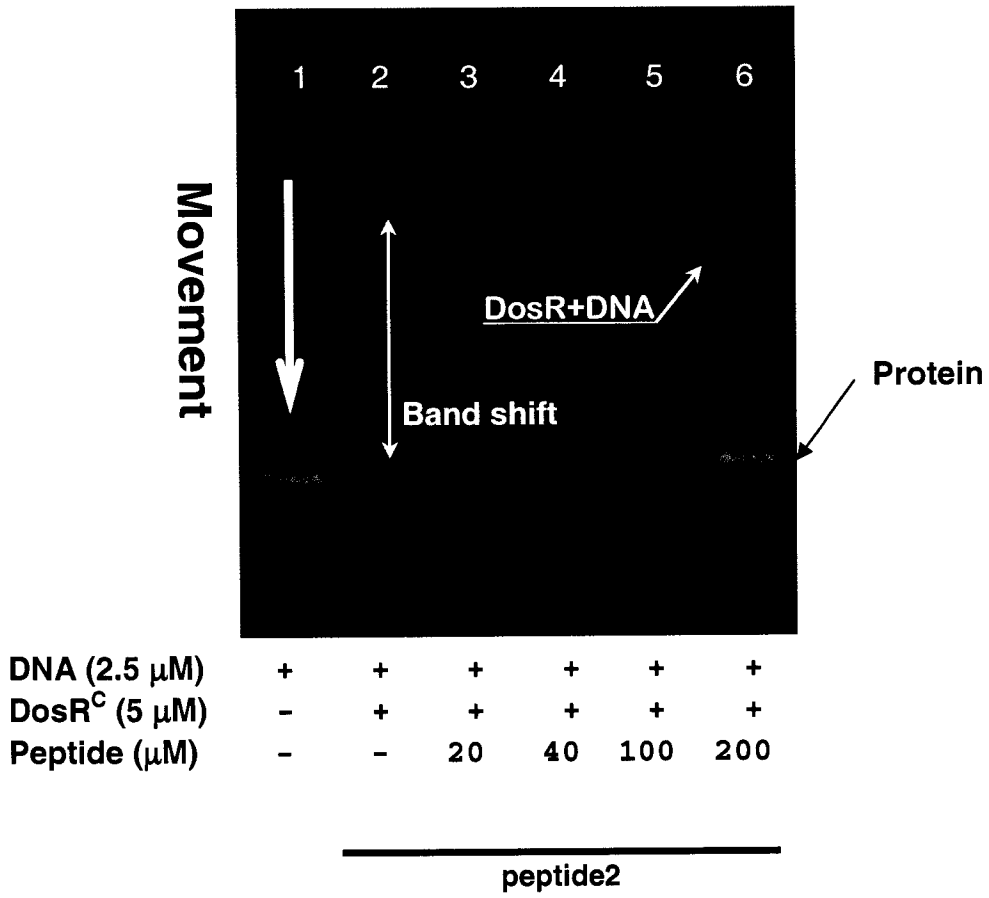


Figure 4.2. A typical pattern for the gel shift assay. Lane 1 is negative control, containing the DNA only. Lane 2 contains both $DosR^C$ and DNA. With the complex formation, the complex band (top) moves a lot slower than free DNA (bottom).

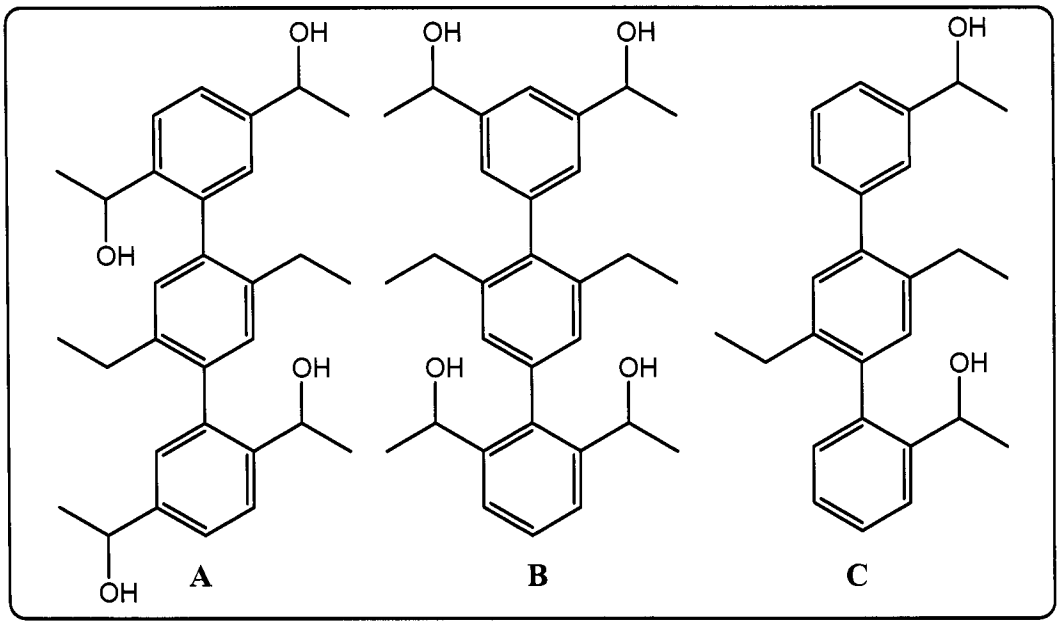


Figure 4.3. The design of terphenyls with two sets of functionalities. In A, the compound has a rotational symmetry; in B the molecule has a C2 symmetry. Considering the sterically demanding nature of A and B, the molecule can be simplified to C for further studies.

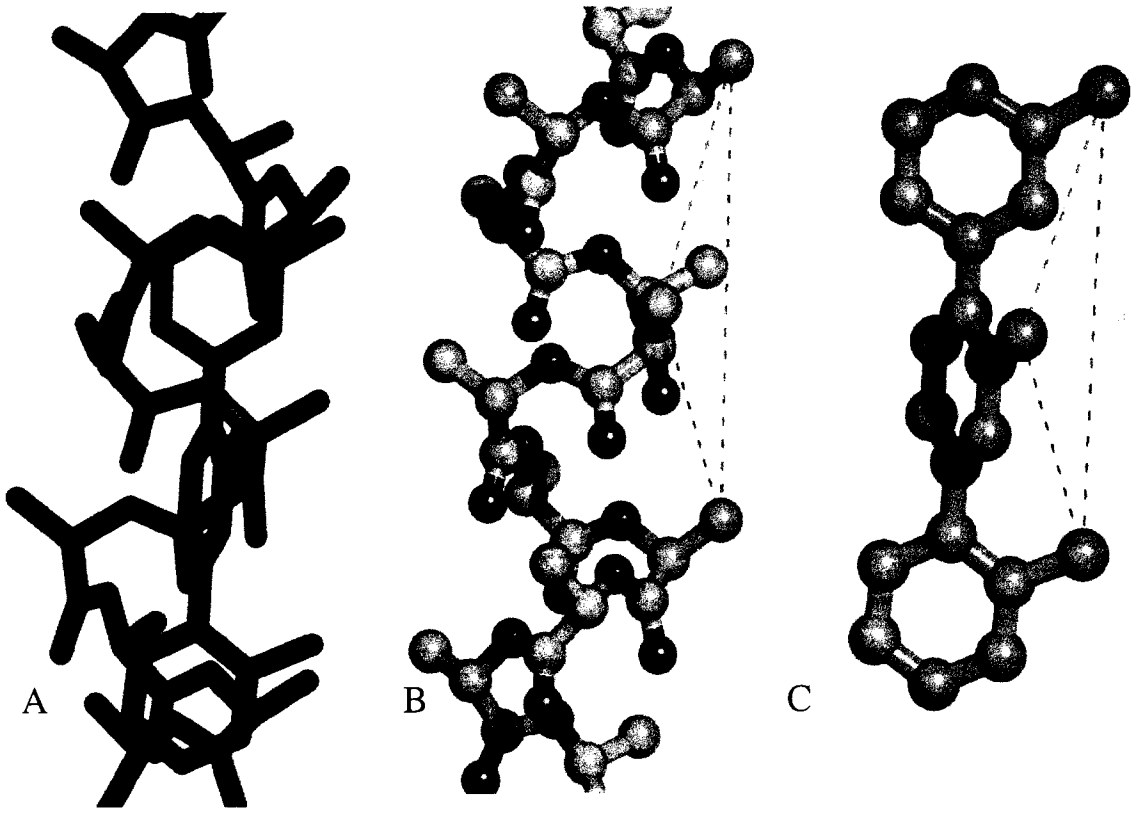


Figure 4.4. 3,2′,2″-trimethyl-p-terphenyl was employed as a model compound for computer modeling. Initial low-energy conformer search results indicated the its lowest energy conformer has torsion angles and distances between substituted groups that are very similar to those residues at *i*, *i*+4, *i*+7 in an α-helix. A is the overlay of low energy conformer with the crystal structure of a polyalanine α-helix. They match well in their presentation of side chain. B and C are schematic representations for the distances between methyl groups at *i*, *i*+4, *i*+7 in that α-helix and 3, 2′, 2″ on the p-terphenyl.

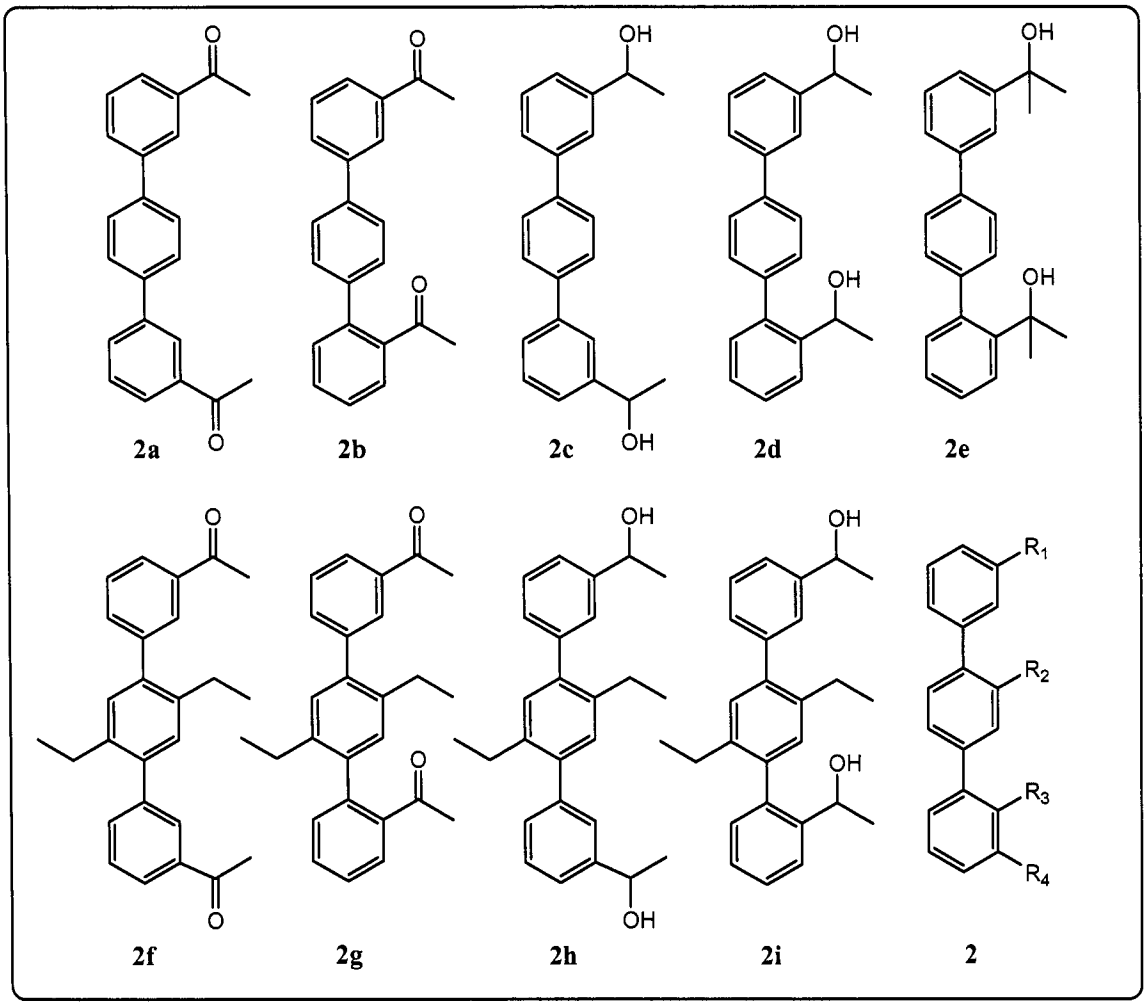


Figure 4.5. A small library of compounds synthesized using Suzuki coupling. Compound **2** is a generic structure showing the substitutions on the terphenyl. The compounds in the second line contain a pair of ethyl substitutions on the middle ring.

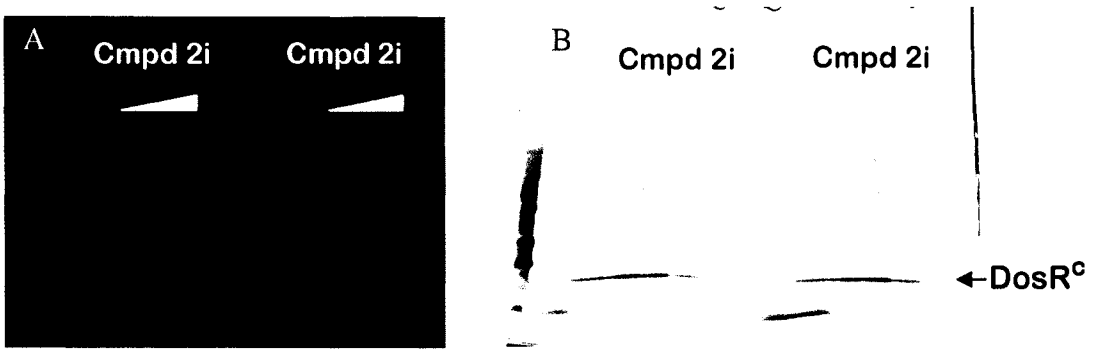


Figure 4.6. Gel shift assay results and SDS-PAGE of compound **2i**. A describes two repeated experiments for compound **2i** in a range from 37.5 μM to 150 μM. B is the SDS-PAGE of the same reaction mixture. The bands at the bottom are $DosR^C$. In B, as the concentration of the inhibitor increased, the band of the $DosR^C$ weakened.

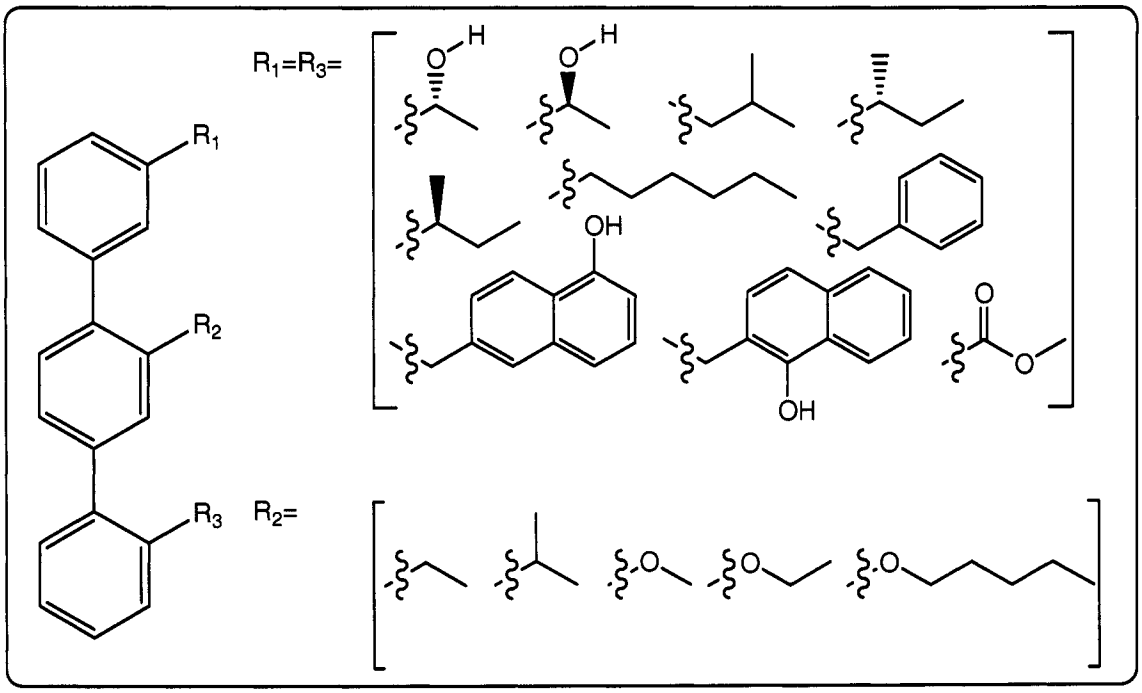


Figure 4.7. The library that was constructed for virtual screening.

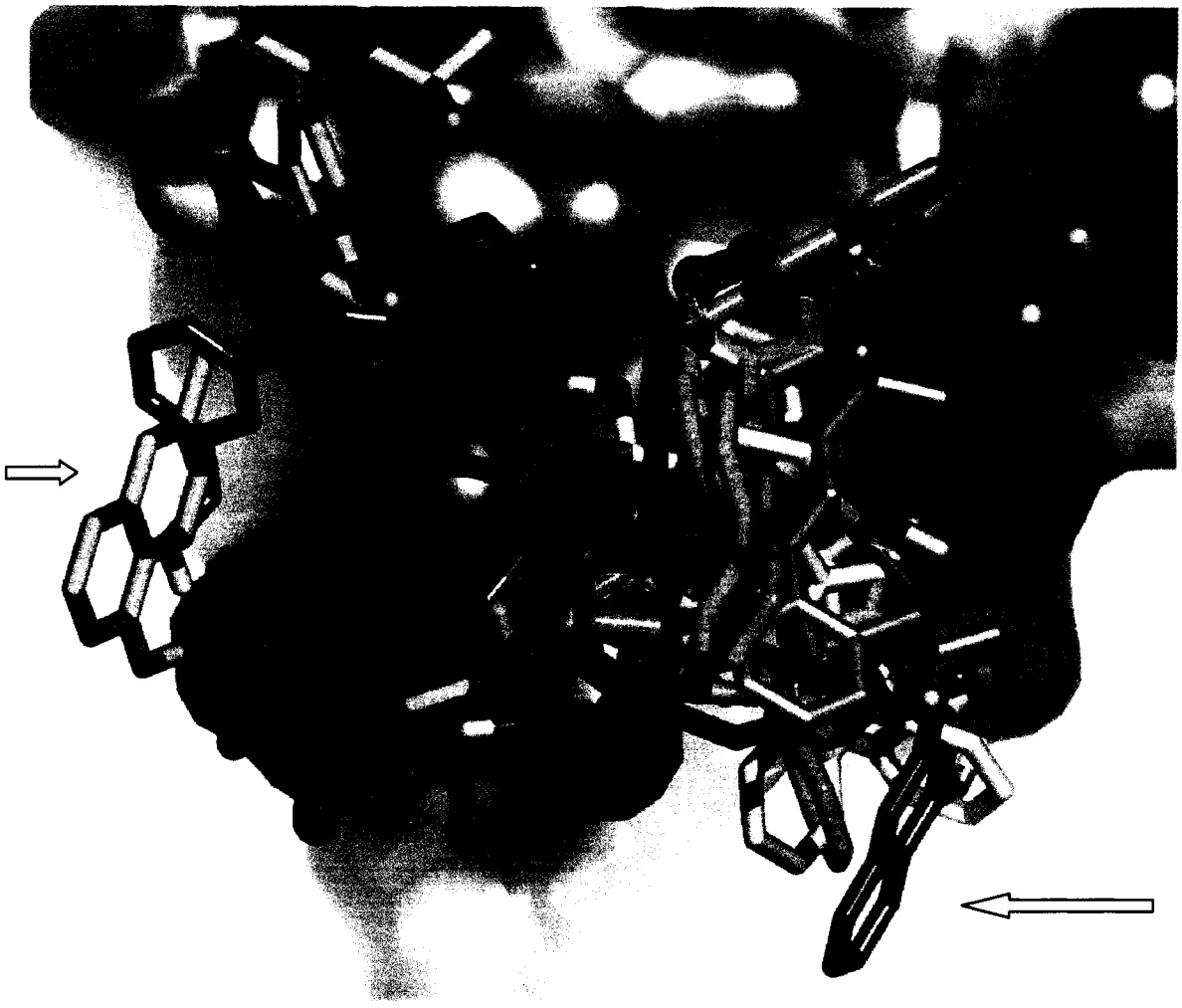


Figure 4.8. A summary of the virtual screening results for the 10×5×10 library. 10 top hits have been superimposed with **3a** (in red), which has been modeled to have an excellent binding affinity towards the DosR interface. Most of the terphenyl compounds are either totally pointing towards the solvent (indicated by the block arrows) or bind at a different pocket. Even for those that can bind nicely for the phenyl ring, they fail to position the second functionality into the right pocket.

Figure 4.9. Surface representation of the inhibitor occupied DosR interface (grey), viewed from the top. The large surface is open to solvent from the top and both ends when DosR is in a monomeric state. The bound molecules depicted here are compounds **3** and **3a**. The superimposition of both molecules suggested a very similar binding mode. Both molecules are adopting a low energy conformation calculated by FLO96. The surface is slightly transparent to show the portions of **3**, **3a** that are less visible from a top view. Nitrogen atoms are in blue, oxygen atom in red and carbon atoms in green or pink.

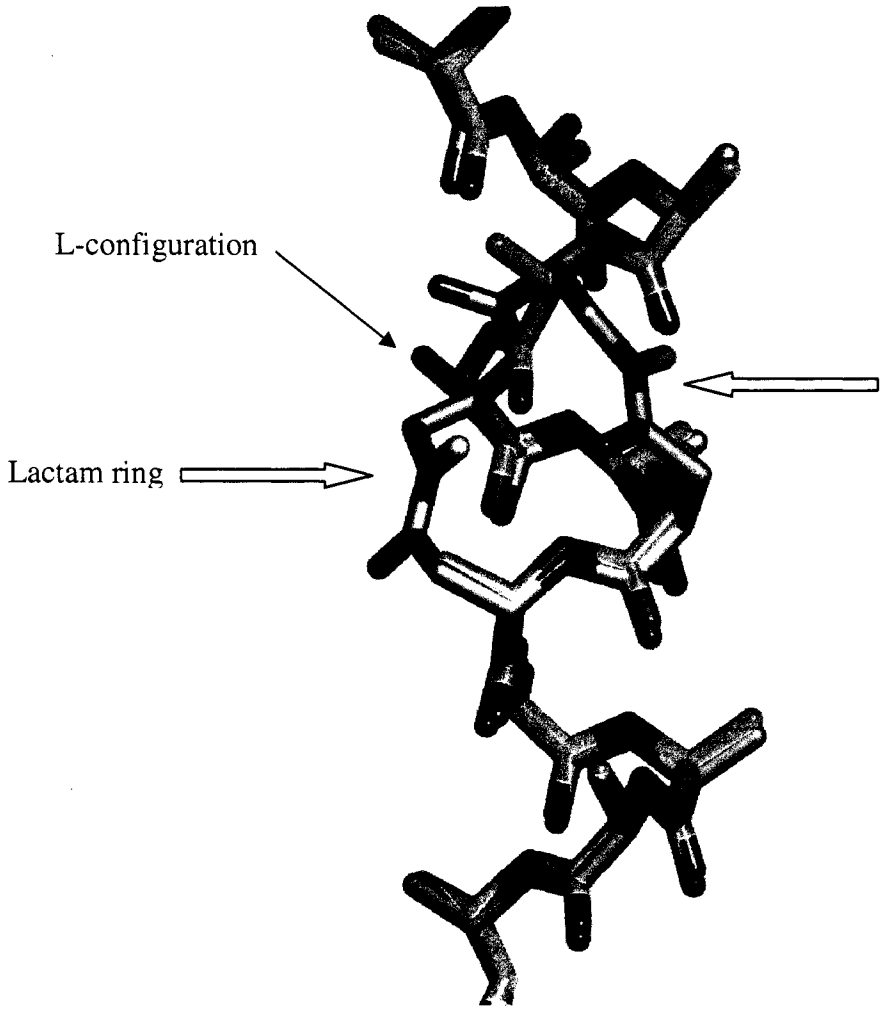


Figure 4.10. Computer modeling results of a doubly cyclized peptide overylaying with a polyalanine α-helix. The conformer shown here has the lowest energy in a FLO96 conformer search. This peptide matches very well with the polyalalinine helix. The block arrows point out the lactam rings formed by side chain cyclization. D-amino acids are incorporated in the peptide and the arrow points out the difference of an L-amino acid as opposed to D amino acid. The polyalanine α-helix is in green and bicyclic peptide is in orange.

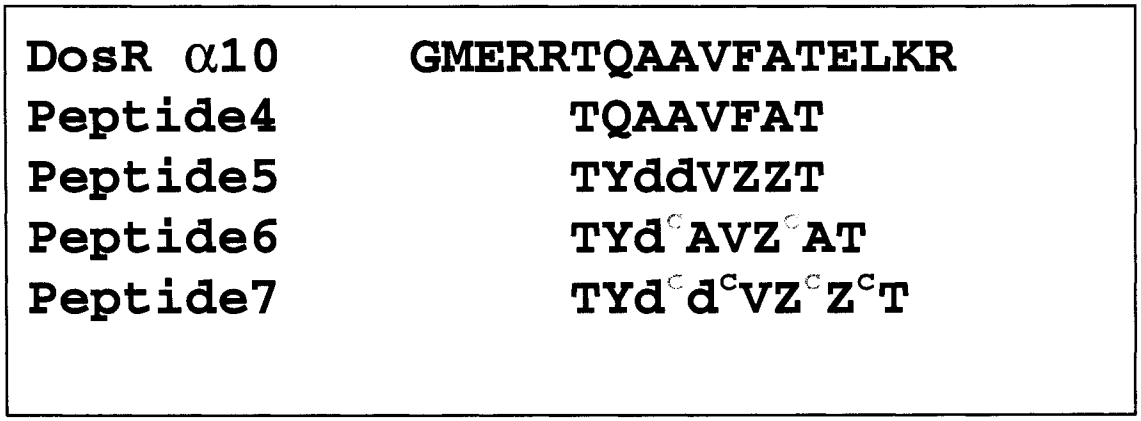


Figure 4.11. The Sequences of peptide **4-7** as compared to α10 of DosR. Z = Dap-OH, c= cylization, the green pair forms one lactam ring, and the blue pair the other. The residues in red are the "TVT" which have been conserved. Amino acids in lower cases are D-amino acids. All the other amino acids are presented in the standard one letter code. All peptides are acetylated at the N-terminal and amidated at the C-terminal.

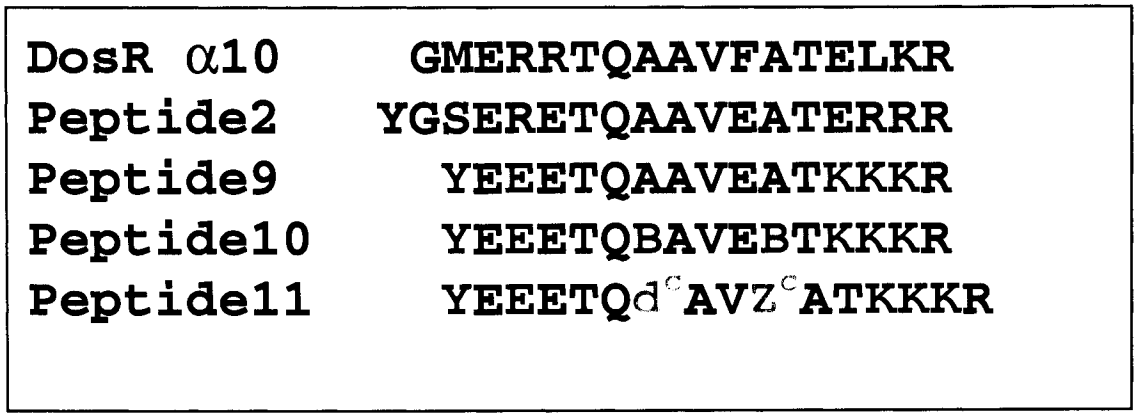


Figure 4.12. The sequences of 16mer peptide **9-11** as compared to peptide **2** and DosR α10 native sequence. B=Aib, the green pair forms one lactam ring. Amino acids in lower cases are D-amino acids. Residues in blue have been mutated from peptide **2**. All peptides are acetylated at the N-terminal and amidated at the C-terminal.

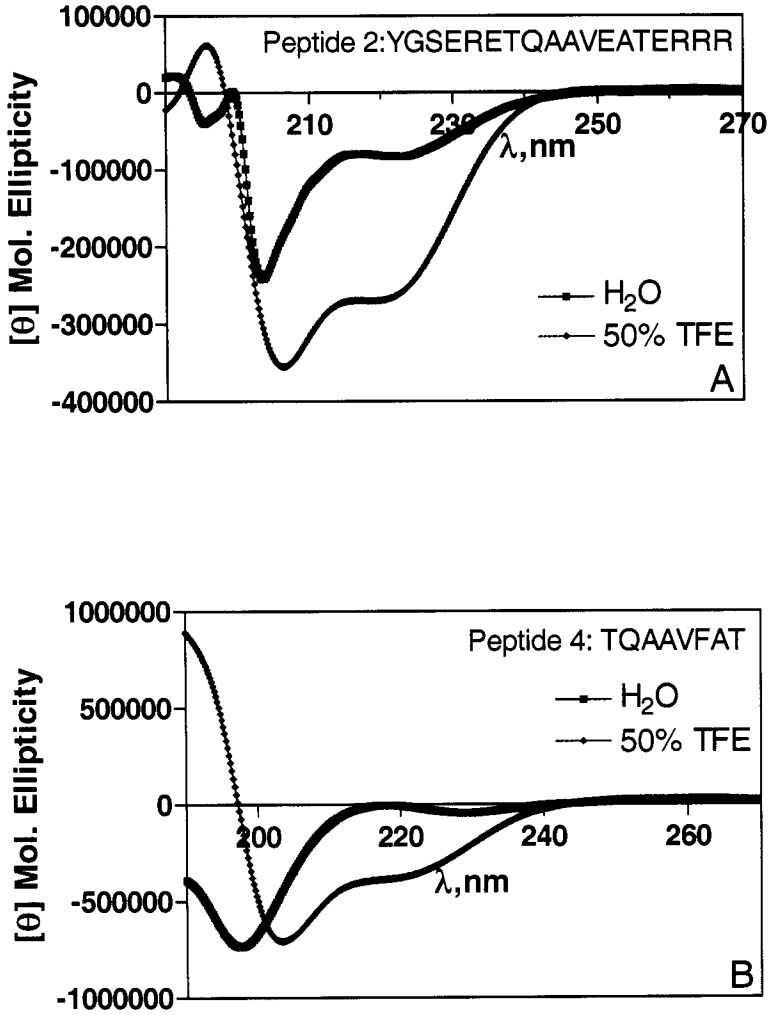


Figure 4.13. Circular dichrism spectra of control peptides in water and 50% TFE. At 20°C, the spectra of 10 μM peptides were recorded at solution between 190-270 nm. The spectra in orange are measured in 50% TFE/H_2O mixture, and the spectra in blue are measured in H_2O. The peptide sequences are also attached on each spectrum. A is for peptide **2**; B is for peptide **4**.

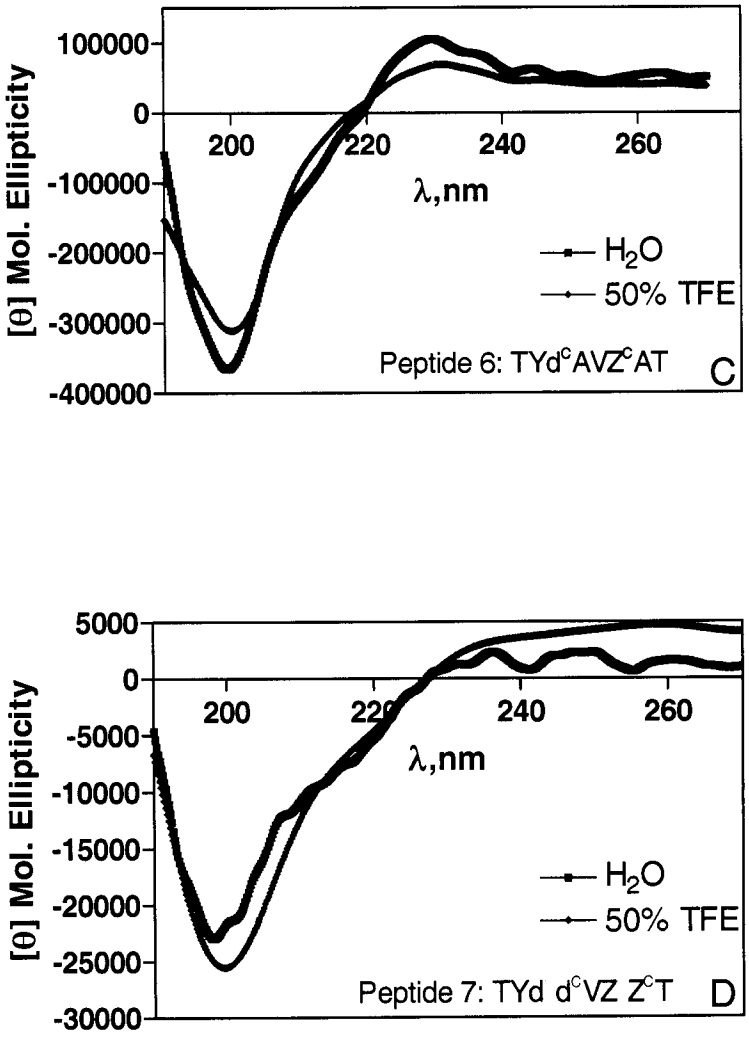


Figure 4.14. Circular dichrism spectra of cyclic peptides in water and 50% TFE. The spectra were taken between 190-270 nm at 20°C. The spectra in orange are measured in 50% TFE/H_2O mixture, and the spectra in blue are measured in H_2O. The peptide sequences are also attached on each spectrum. C is for monocyclic peptide **6**; D is for bicyclic peptide **7**.

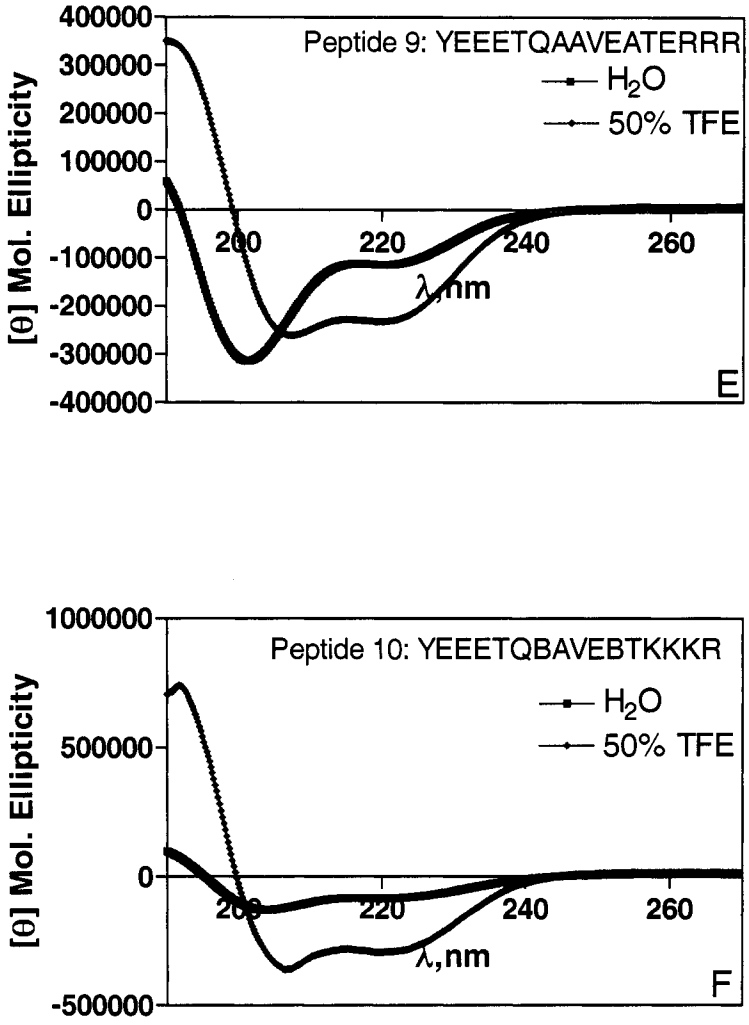


Figure 4.15. Circular dichrism spectra of 16mer linear peptides in water and 50% TFE. The spectra were taken between 190-270 nm at 20°C. The spectra in orange are measured in 50% TFE/H_2O mixture, and the spectra in blue are measured in H_2O. In both cases, a significant increase in helicity was observed in 50% TFE. E is for peptide **9**; F is for peptide **10**.

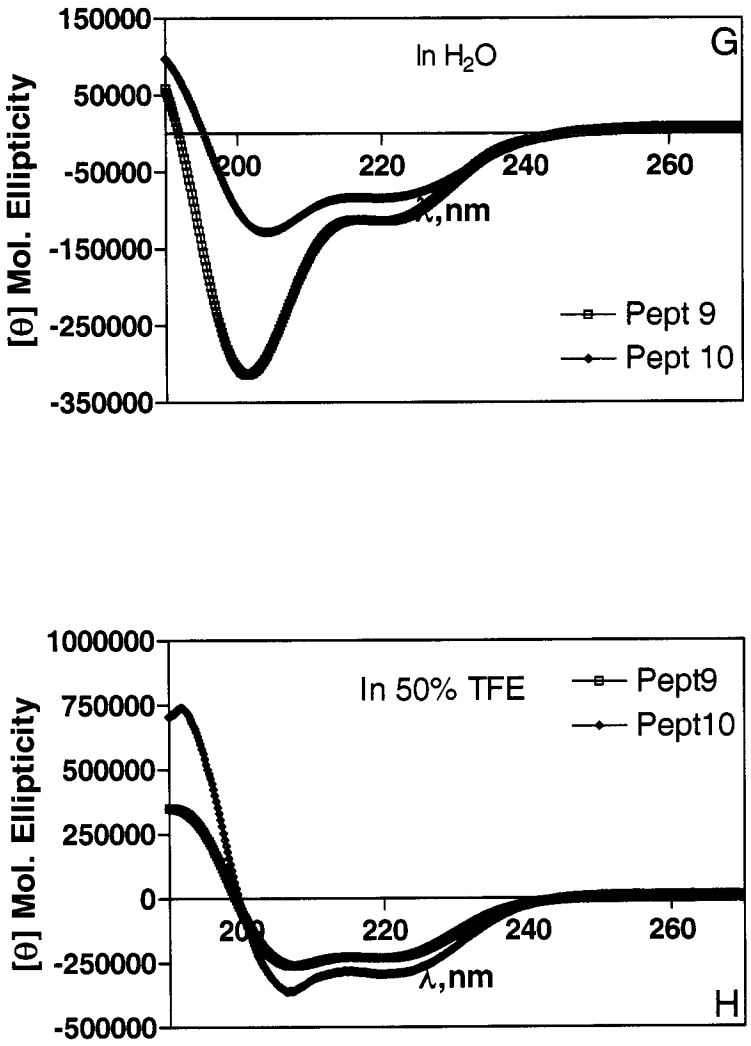


Figure 4.16. Overlay of of CD spectra of 18mer linear peptides in water and 50% TFE. The spectra in pink are for peptide **10** and blue for peptide **9**. The spectra measured in H_2O are in G; the spectra measured in 50% TFE are in H.

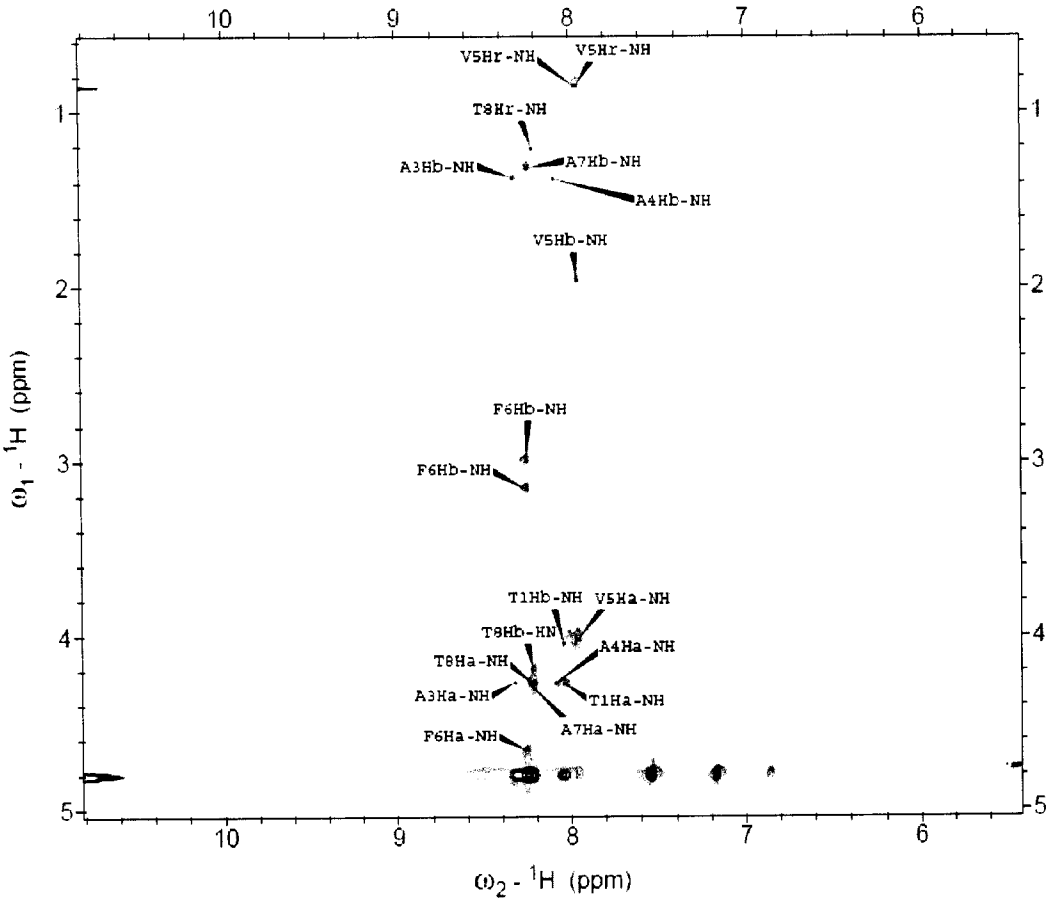


Figure 4.17. Fingerprint region of 2D-TOCSY of peptide **4**. This spectrum was analyzed with combined information from 2D-ROESY. Only the H^{α}-H^{N} coupling region is shown here for clarification purposes. Also refer to **Figure 4.18** for chemical shift deviation (CSD) analysis.

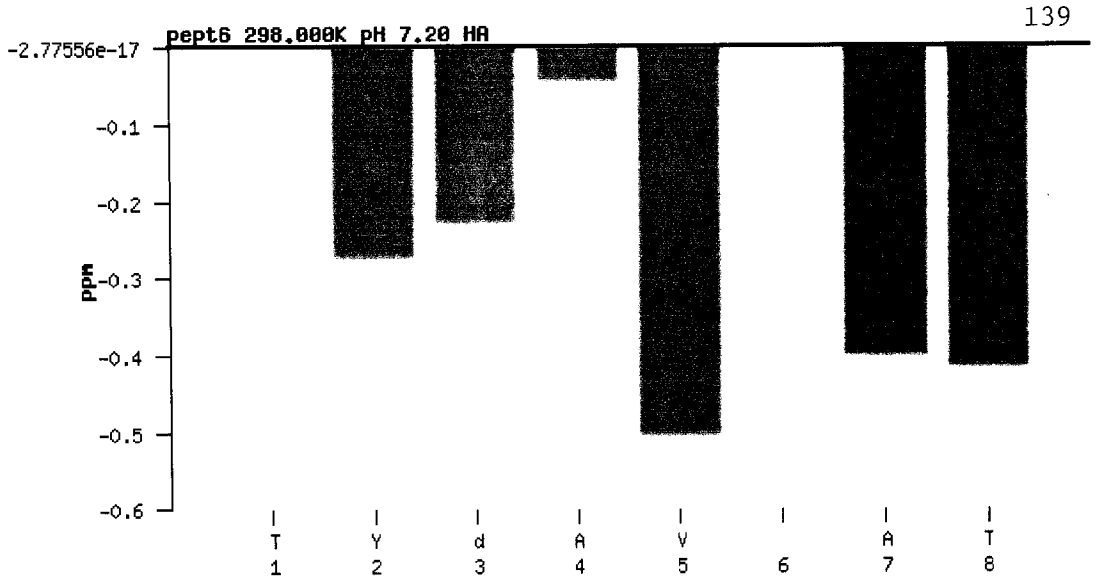


Figure 4.18. Chemical shift deviation (CSD) of Hα for peptide **6**. CSD is consistently negative, characteristic of an α-helical structure. Residue 6 is a D-Asp. Since there are no chemical shift data available for D-amino acids in a random coil, it was left blank. The figure was produced by using the software CSDb from Prof. Andersen's website at the University of Washington, department of chemistry.

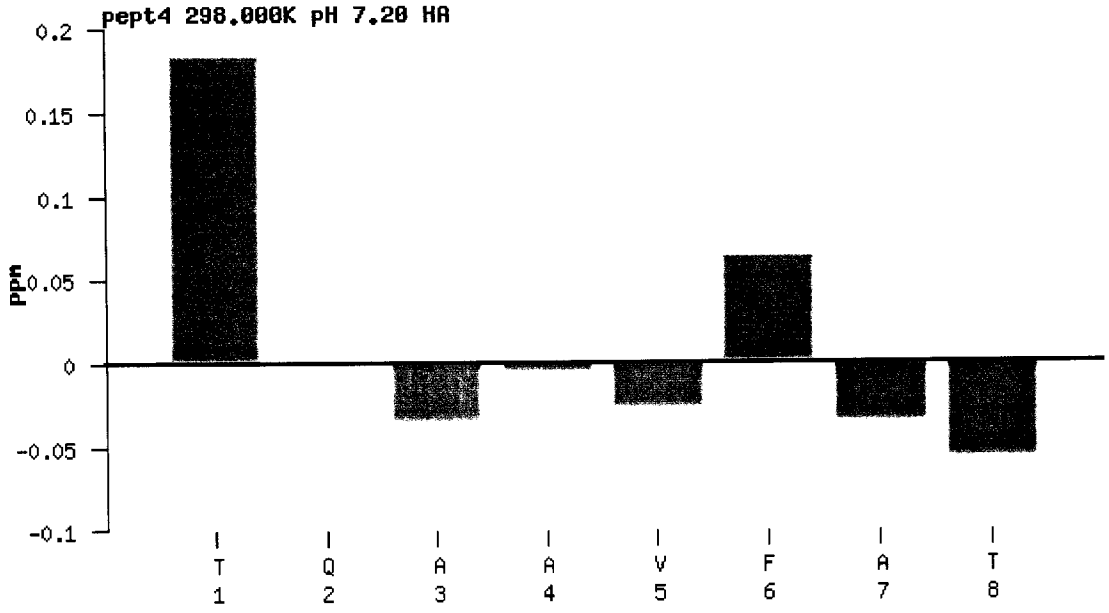


Figure 4.19. Chemical shift deviation (CSD) of Hα for peptide **4**. CSD is defined as $\Delta\delta = \delta - \delta_{ref}$, where δ_{ref} is based on statistical analysis of the random coil. The figure was produced by using the software CSDb from Prof. Andersen's website at the University of Washington, department of chemistry.

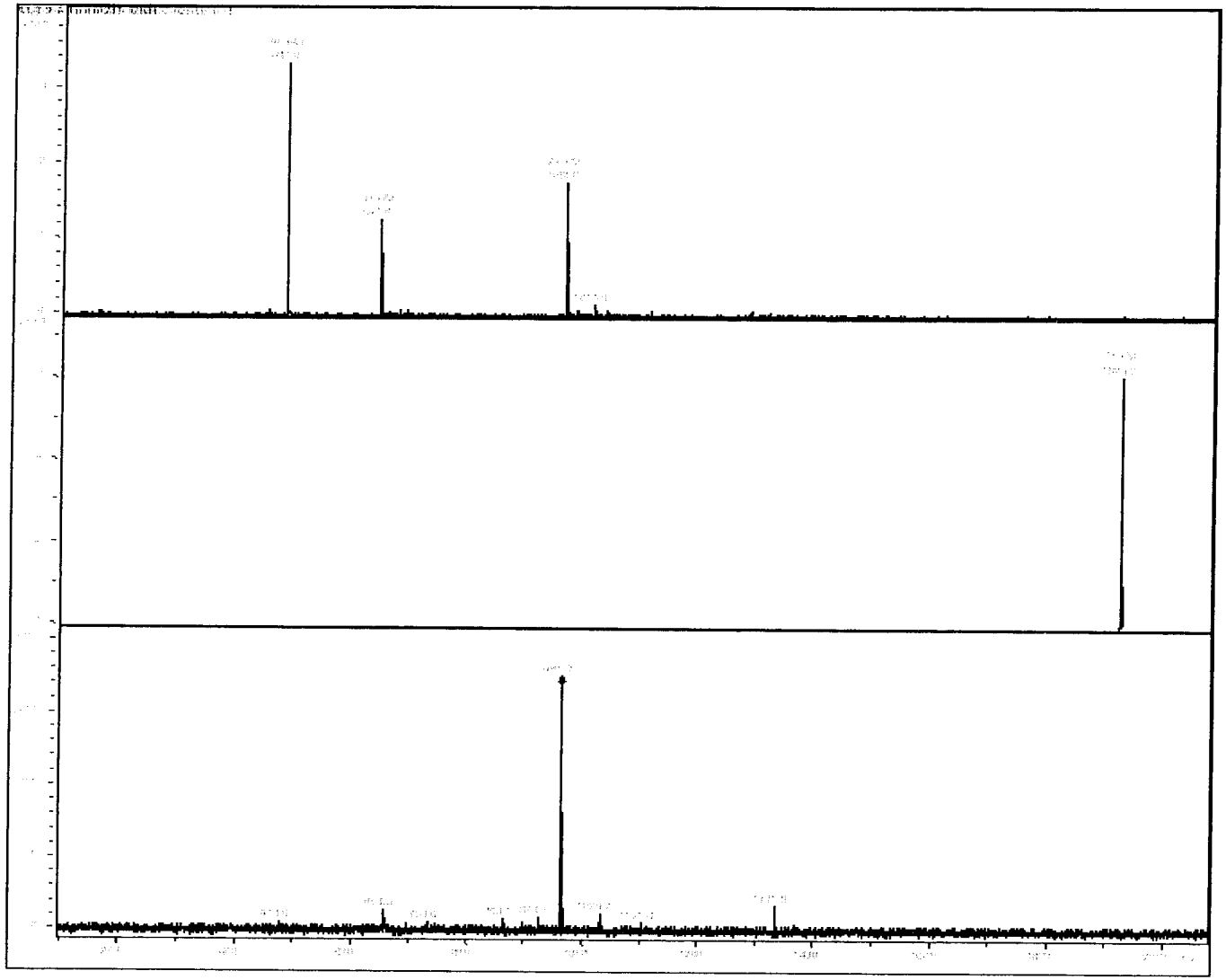
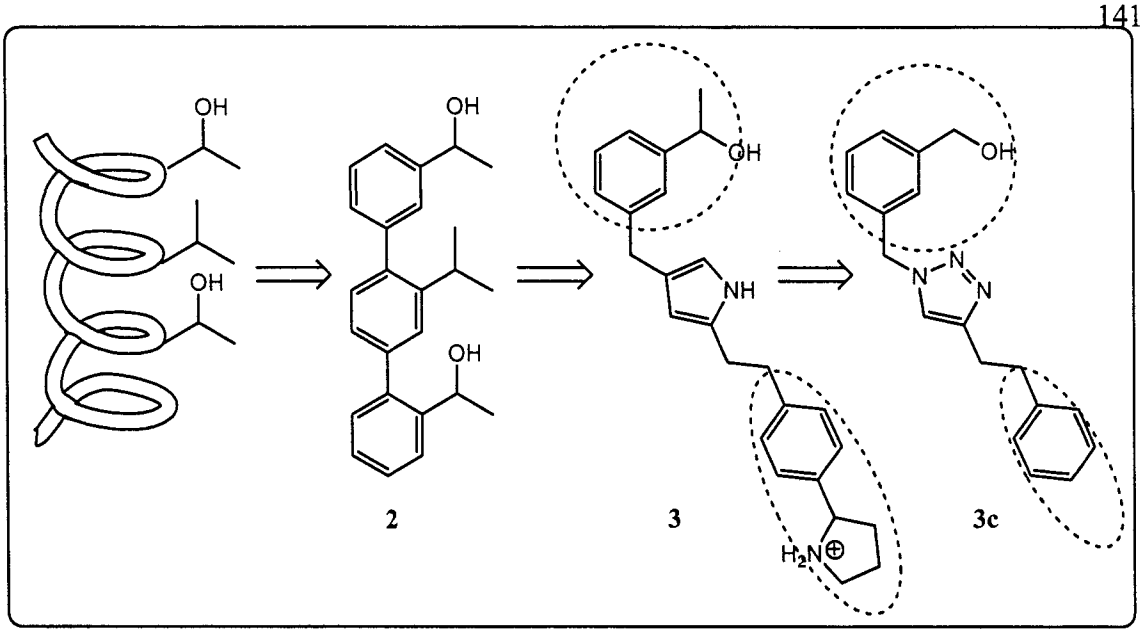
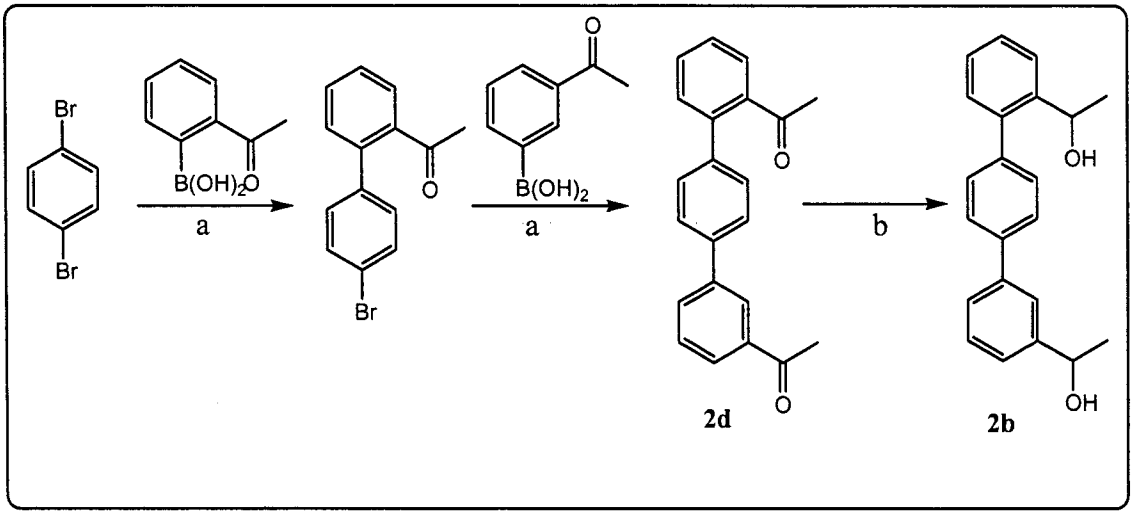


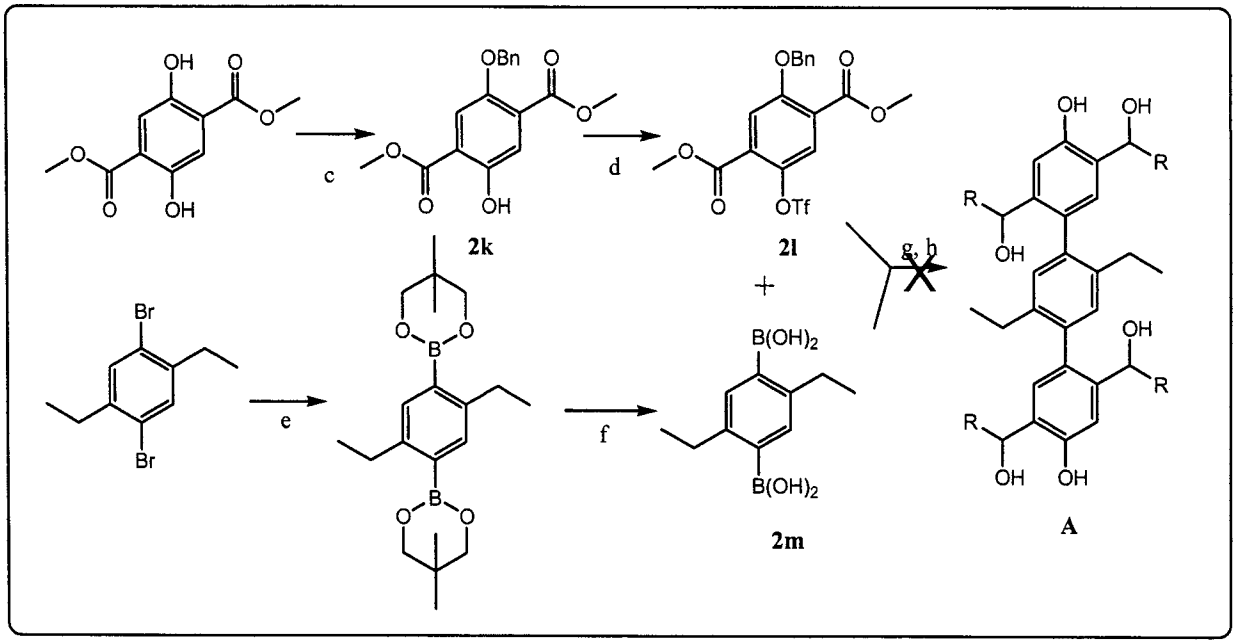
Figure 4.20. Mass spectrum of peptide **10**. The $(M+2H)^{2+}$, $(M+3H)^{3+}$, $(M+4H)^{4+}$ peaks are clearly observed and $(M+H)^{+}$ was calculated from deconvolution.



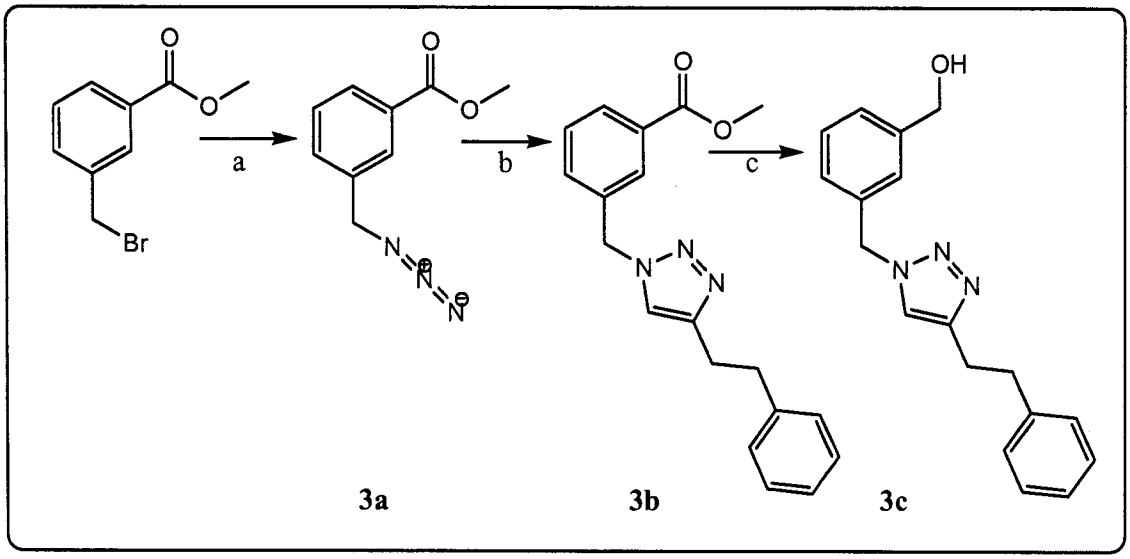
Scheme 4.1. Schematic representation of the progress of small molecule inhibitor design. Firstly, side chain residues TVT can be symbolized by 1-hydroxyethyl, isopropyl and 1-hydroxymethyl respectively. This model (**2**) was optimized to benzylpyrrole derivative for potentially higher binding affinity. Further modification led to compound **3c**, which gave synthetic ease and opportunities for variation. The portion in the dotted circle is preserved because of its favorite interaction with DosR monomer.



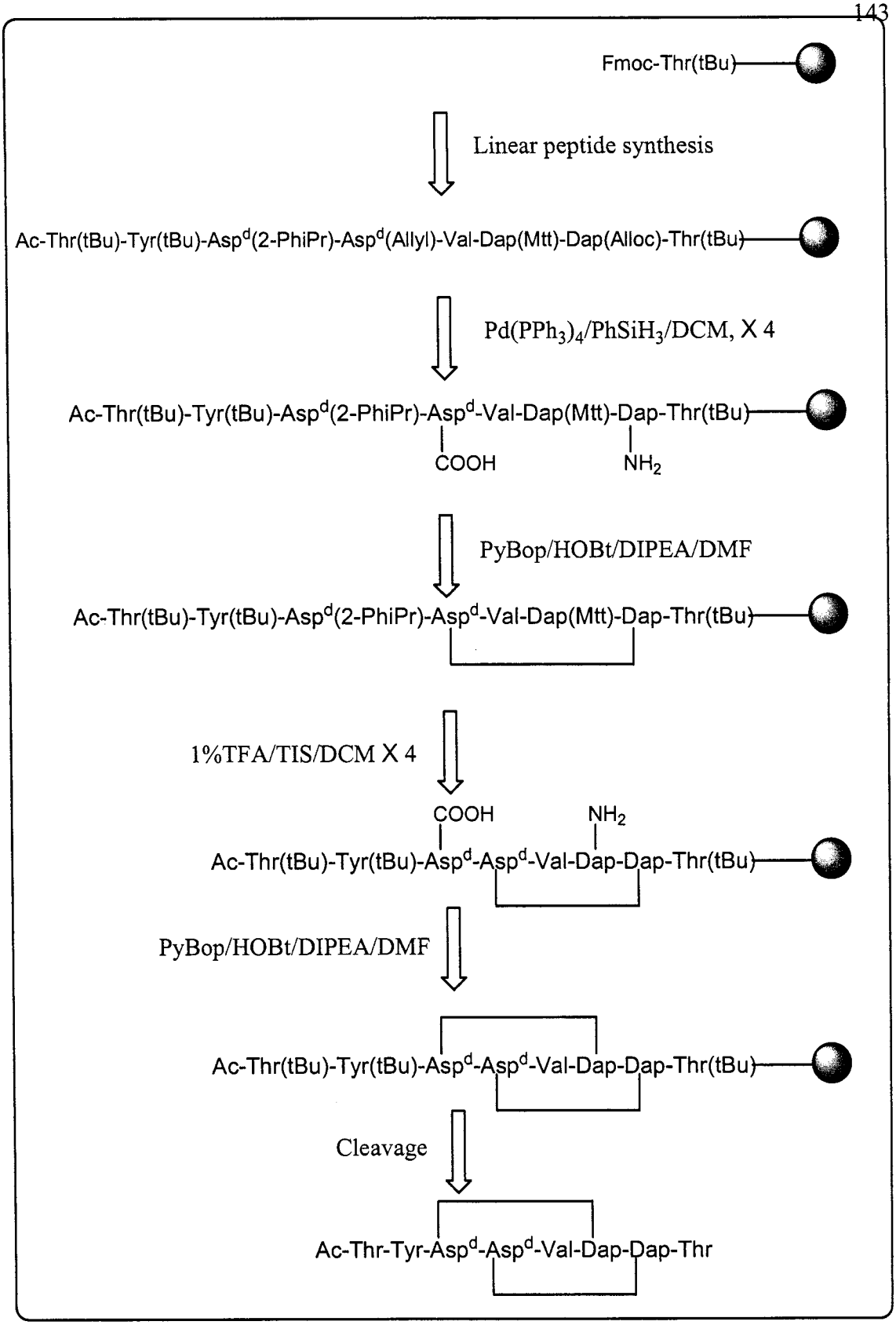
Scheme 4.2. Synthesis of terphenyl model compound through sequencial Suzuki coupling. a) 1:1 molar ratio, 1-5% $Pd(PPh_3)_4$, DME/EtOH=9:1 volume ratio, Na_2CO_3(2M), 85°C, microwave, 5 min. b) $NaBH_4/LiAlH_4$, 10 eq., CH_2Cl_2, RT 2 hr. DME=1,1-dimethoxyethane.



Scheme 4.3. A proposed synthetic route for the synthesis of functional terphenyls. c) BnBr, NaH, THF, RT, 3 hr; d) Tf_2O, Pyridine, N_2, 0°C to RT overnight; e) bis(neopentyl glycolato)dibron, $PdCl_2(dppf)$, KOAc, DMSO, 80°C, 4 hr; f) 1M HCl, RT, 4 hr; g) $Pd(OAc)_2$ 1-5%, bornic acid, $K_3PO_4 \cdot H_2O$, THF, L= 2-(2',4',6'-triisopropylbiphenyl)diphenyl phosphine, L:Pd=3:1 RT 24 hr; h) 1)MgBrR, Et_2O, - 45°C, 4 hr ii) 2M HCl aq., RT, 4 hr.



Scheme 4.4. Synthesis of compound **3c** with "click chemistry". a) NaN_3, DMF, RT, 2 hr. b) 3-phenylpropyne, sodium L-ascorbate, $CuSO_4$, μ-wave 10 min. c) $NaBH_4/LiAlH_4$,CH_2Cl_2.



Scheme 4.5. Solid phase synthesis for bicyclic peptide **7**.

Table 4.1. Compounds in the terphenyl library and their IC50s.
* IC50 measured from a gel shift binding assay.
a. The assay was performed with 12.5-37.5% DMSO in the buffer.
b. The assay was performed with 40% DMSO in the buffer.

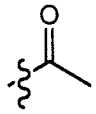
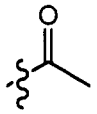
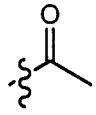
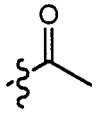
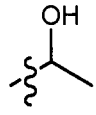
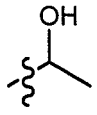
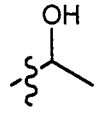
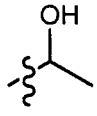
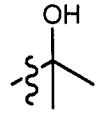
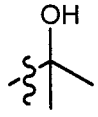
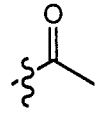
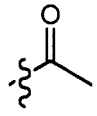
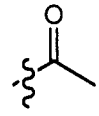
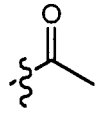
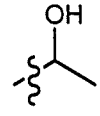
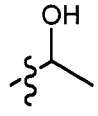
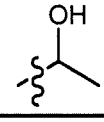
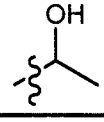
Compoud	R1	R2	R3	R4	IC50*
2a		H	H		~60 μM[a]
2b		H		H	~15 μM[b]
2c		H	H		~250 μM[b]
2d		H		H	>500 μM[b]
2e		H		H	–
2f			H		~250 μM[b]
2g				H	>250 μM[b]
2h			H		~750 μM[b]
2i				H	~40 μM[b]

Table 4.2. Molar Ellipticities $[\theta]$ (deg cm^2 dmol^{-1}) at λ= 222, 203 nm, Ellipticity ratios $\theta_{222}/\theta_{203}$, and relative helicities for 8mer peptides **4**, **6**, **7** in H_2O and 50% TFE. All the ellipticity values were measured at 20°C.

a. Relative helicity is calculated based on $([\theta]_{222})_x/([\theta]_{222})_{4\ (50\%\ TFE)}$.
b. For cyclic peptides, due to the existence of D-amino acid, the CD signal is dramatically offset. In this case, only change in relative helicity is meaningful.

Peptide	$[\theta]_{222}$	$[\theta]_{203}$	$\theta_{222}/\theta_{203}$	Solvent	Relative Helicity[a]
4	-2988	-87500	0.034	H_2O	0.06
4	-52142	-96428	0.541	50% TFE	1.00
6	9667	-19833	-0.487	H_2O	-0.19[b]
6	7342	-21583	-0.340	50% TFE	-0.14[b]
7	-83.3	-767	0.108	H_2O	0.001[b]
7	-73.3	-717	0.102	50% TFE	0.001[b]

Table 4.3. Molar Ellipticities $[\theta]$ (deg cm^2 dmol^{-1}) at λ= 222, 208 nm, ellipticity ratios $\theta_{222}/\theta_{208}$, and relative helicities for peptides **2**, **9**, **10** in H_2O and 50% TFE.

a. Relative helicity is calculated based on $([\theta]_{222})_x/([\theta]_{222})_{\mathbf{10}\ \text{in 50\% TFE}}$. All the ellipticity values were measured at 20°C.

Peptide	$[\theta]_{222}$	$[\theta]_{208}$	$\theta_{222}/\theta_{208}$	Solvent	Relative Helicity[a]
2	-4575	-13420	0.341	H_2O	0.25
2	-14600	-18917	0.772	50% TFE	0.81
9	-7000	-12666	0.553	H_2O	0.63
9	-14224	-16230	0.876	50% TFE	0.79
10	-5130	-710	0.737	H_2O	0.28
10	-18092	-21773	0.831	50% TFE	1.00

Notes to Chapter 4

1. Liu J. Natural terphenyls: developments since 1877. *Chem. Rev.* **2006**, 106, 2209-2223.

2. Maeyama K., Yonezawa N. Synthesis of substituted m-terphenyls. *Recent Res. Dev. Org. Chem.* **2003**, 7, 63-69.

3. Maeyama K., Yonezawa N. Synthesis of substituted o- terphenyl via electrophilic aromatic substitution and other routes. *Recent Res. Dev. Org. Chem.* **2003**, 7, 53-61.

4. Greenfield A.A., Butera J. A., Caufield C. E. Convenient synthesis of functionalized terphenyls. *Tet. Lett.* **2003**, 44, 2729-2732.

5. Uchida K., Segawa T., Takahashi Y. Absorption bands in the absorption edge region of p-terphenyl crystals. *J. Lumin.* **1994**, 60-61, 795-797.

6. Minoru A. *Spectrochimica Acta, Part A* **1984**, 40, 367-371.

7. Yin H., Hamilton A. D. Strategies for targeting protein-protein interactions with synthetic agents. *Angew. Chem. Int. Ed. Engl.* **2005**, 27, 4130-4163.

8. Kutzki O., Park H. S., Ernst J. T., Orner B. P., Yin H., Hamilton A. D. Development of a potent Bcl-x(L) antagonist based on alpha-helix mimicry. *J. Am. Chem. Soc.* **2002**, 40, 11838-11839.

9. Orner B. P., Ernst J. T., Hamilton A. D. Toward proteomimetics: terphenyl derivatives as structural and functional mimics of extended regions of an alpha-helix. *J. Am. Chem. Soc.* **2001**, 123, 5382-5383.

10. Yin H., Lee G. I., Park H. S., Payne G. A., Rodriguez J. M., Sebti S. M., Hamilton A. D. Terphenyl-based helical mimetics that disrupt the p53/HDM2 interaction. *Angew. Chem. Int. Ed. Engl.* **2005**, 18, 2704-2707.

11. Ernst J. T., Kutzki O., Debnath A. K., Jiang S., Lu H., Hamilton A. D. Design of a protein surface antagonist based on alpha-helix mimicry: inhibition of gp41 assembly and viral fusion. *Angew. Chem. Int. Ed. Engl.* **2002**, 41, 278-281.

12. Ishiyama T., Murata M., Miyaura N. Palladium (0)-catalyzed cross-coupling reaction of alkoxydiboron with haloarenes: a direct procedure for arylboronic esters. *J. Org. Chem.* **1995**, 60, 7508-7510.

13. Ni W., Fang H., Springsteen G., Wang B. The design of boronic acid spectroscopic reporter compounds by taking advantage of the pK_a-lowering effect of diol binding: nitrophenol-based color reporters for diols. *J. Org. Chem.* **2004**, 69, 1999-2007.

14. Singh R. Viciu M. S., Kramareva N., Navarro O., Nolan S. P. Simple (imidazol-2-ylidene)-Pd-acetate complexes as effective prcatalysts fpr sterically hindered Susuki-Miyaura couplings. *Org. Lett.* **2005**, 7, 1829-1832.

15. DeVasher R. B., Moore L. R., Shaughnessy K. H. Aqueous-phase, palladium-catalyzed cross-coupling of aryl bromides under mild conditions, using water-soluble, sterically demanding alkylphosphines. *J. Org. Chem.* **2004**, 69, 7919-7927.

16. Barder T. E., Walker S. D., Martinelli J. R., Buchwald S. L. Catalysts for Suzuki-Miyaura coupling processes: scope and studies of the effect of ligand structure. *J. Am. Chem. Soc.* **2005**, 127, 4685-4696.

17. Fang H., Kaur G., Yan J., Wang B. An efficient synthesis of sterically hindered arylboronic acids. *Tet. Lett.* **2005**, 46, 1671-1674.

18. Taylor R. D., Jewsbury P. J., Essex J. W. A review of protein-small molecule docking methods. *J. Comput. Aid. Mol. Design* **2002**, 16, 151–166.

19. ThistleSoft, Inc. FLO96, ThistleSoft, Colebrook, CT, **1996**.

20. Kolb H. C., Sharpless K. B. The growing impact of click chemistry on drug discovery. *Drug Discovery Today* **2003**, 8, 1128-1137.

21. Demko Z. P., Sharpless K. B. A click chemistry approach to tetrazoles by Huisgen 1,3-dipolar cycloaddition: synthesis of 5-acyltetrazoles from azides and acyl cyanides. *Angew. Chem. Int. Ed. Engl.* **2002**, 41, 2113-2116.

22. Kolb H. C., Finn M. G., Sharpless K. B. Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem. Int. Ed. Engl.* **2001**, 40, 2004-2021.

23. Horal P., Svennerholm B., Jeansson S., Rymo L., Hall W.W. & Vahlne A. Continuous epitopes of the human immunodeficiency virus type 1 (HIV-1) transmembrane glycoprotein and reactivity of human sera to synthetic peptides representing various HIV-1 isolates. *J. Virol.* **1991**, 65, 2718–2723.

24. Barlow D. J. Peptide drug design. *Smith and Williams' Introduction to the principles of drug design and action (4th Ed.)* **2006**, 327-354.

25. Mizejewski G. J. Peptides as receptor ligand drugs and their relationship to G-coupled signal transduction. *Expert Opin. Invest. Drugs* **2001**, 10, 1063-1073.

26. Hamman J. H., Enslin G. M., Kotze A. F. Oral delivery of peptide drugs: barriers and developments. *BioDrugs* **2005**, 19, 165-177.

27. Kotha S., Lahiri K. Post-assembly peptide modifications by chemical methods. *Curr. Med. Chem.* **2005**, 12, 849-875.

28. Kastantin M. and Tirrell M. Entropic stabilization of alpha-helix formation in peptide amphiphiles. *Poster presented at the MROP conference*, 31 January, **2005**.

29. Hol W.G. The role of the α-helix dipole in protein structure and function. *Prog. Biophys. Mol. Biol.* **1985**, 45, 149-195.

30. Kemp D. S., Allen T. J., Oslick S. L. The energetics of helix formation by short templated peptides in aqueous solution. 1. Characterization of the reporting helical template Ac-Hel1. *J. Am. Chem. Soc.* **1995**, 117, 6641-6657.

31. Jackson D. Y., King D. S., Chmielewski J., Singh S., Schultz P. G. General approach to the synthesis of short .alpha.-helical peptides. *J. Am. Chem. Soc.* **1991**, 24, 9391-9392.

32. Taylor J. W. The synthesis and study of side-chain lactam-bridged peptides. *Biopolymers* **2002**, 1, 49-75.

33. Schmiedeberg N., Kessler H. Reversible backbone protection enables combinatorial solid-phase ring-closing metathesis reaction (RCM) in peptides. *Org. Lett.* **2002**, 4, 59 -62.

34. Cabezas E., Satterthwait A. C. The hydrogen bond mimic approach: solid-phase synthesis of a peptide stabilized as an α-helix with a hydrazone link. *J. Am. Chem. Soc.* **1999**, 121, 3862-3875.

35. Maeda M., Melnyk R. A., Partridge A. W., Liu L. P., Deber C. M.. Transmembrane segment peptides with double D-amino acid replacements: helicity, hydrophobicity, and antimicrobial activity. *Biopolymers* **2003**, 71, 77-84.

36. Cuthbertson A., Indrevoll B. Regioseletive formation, using orthogonal cysteine protection, of an α-conotoxin dimer peptide containing four disulfide bonds. *Org. Lett.* **2003**, 5, 2955-2957.

37. Albericio F. Orthogonal protecting groups for N^{α}-amino and C-terminal carboxy functions in solid-phase peptide synthesis. *Biopolymers* **2000**, 55, 123-139.

38. Chakrabartty A., Doig A. J., Baldwin R. L. Helix capping propensities in peptides parallel those in proteins. *Proc. Natl. Acad. Sci. USA* **1993**, 90, 11332-11336.

39. Banerjee R., Basu G., Chene P., Roy S. Aib-based peptide backbone as scaffolds for helical peptide mimics. *J. Peptide Res.* **2002**, 60, 88–94.

40. Bellanda M., Peggion E., Burgi R., van Gunsteren W., Mammi S. Conformational study of an Aib-rich peptide in DMSO by NMR. *J. Peptide Res.* **2001**, 57, 97-106.

41. Aravinda S.., Shamala N., Roy R. S., Balaram P. Non-protein amino acids in peptide design. *Proc. Indian Acad. Sci.* **2003**, 115, 373-400.

42. Berg J. M., Tymoczko J. L., Stryler L., Clarke N. D. *Biochemistry (5th Ed.)* W. H. Freeman and Company, New York, **2002**.

43. Shepherd N., Hoang H. N., Abbenante G., Fairlie D. P. Single turn peptide alpha helices with exceptional stability in water. *J. Am. Chem. Soc.* **2005**, 127, 2974-2983.

44. Chen Y. H., Yang T. Y., Chau K. H. Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry*, **1974**, 13, 3350-3359.

45. Luo P., Baldwin L. R. Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* **1997**, 36, 8413-8421.

46. Luo P., Braddock D. T., Subramanian R. M., Meredith S. C., Lynn D. G. Structural and thermal dynamic characterization of a bioactive peptide model of apolipoprotein E: side-chain lactam bridges to constrain the conformation. *Biochemistry* **1994**, 33, 12367-12377.

47. Scolnik Y., Portnaya I., Cogan U., Tal S., Haimovitz R., Fridkin M., Elitzur A. C., Deamer D. W., Shinitzky M. Subtle differences in structural transitions between poly-L- and poly-D-amino acids of equal length in water. *Phys. Chem. Chem. Phys.* **2006**, 8, 333-339.

48. Wallimann P., Kennedy R. J., Kemp D. S. Large circular dichroism ellipticities for N-templated helical polypeptides are inconsistent with currently accepted helicity algorithms. *Angew. Chem. Int. Ed. Engl.* **1999**, 38, 1290-1292.

49. Bracken C., Gulyras J., Taylor J. W., Baum J. Synthesis and nuclear magnetic resonance structure determination of an α-helix, bicyclic, lactam-bridged hexapeptide. *J. Am. Chem. Soc.* **1994**, 116, 6431-6432.

50. Wuthrich K. NMR of proteins and nucleic acids. John Wiley and Sons Inc., USA, **1986**.

51. Wishart D. S., Sykes B. D., Richards F. M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* **1991**, 222, 311-333.

52. Wishart D. S., Sykes B. D., Richards F. M. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **1992**, 31, 1647-1651.

53. Wishart D. S., Sykes B. D., Richards F. M. Simple techniques for the quantification of protein secondary structure by 1H NMR spectroscopy. *FEBS Lett.* **1991**, 293, 72-80.

54. Sun Y., Ewing T. J. A., Skillman A. G., Kuntz I. D. CombiDOCK: structure-based combinatorial docking and library design. *J. Comput.-Aided Mol. Des.* 1998, 12, 594

55. Bax A., Davis D. G. *J. Mag. Reson.* **1985**, 63, 207-213.

56. Hwang T. L., Shaka A. J. Cross relaxation without TOCSY: transverse rotating-frame overhauser effect spectroscopy. *J. Am. Chem. Soc.* **1992**, 114, 3157-3159.

57. Piotto M., Saudek V., Sklenár V. A simple modification to the WATERGATE water suppression scheme. *J. Biomol. NMR* **1992**, 2, 661–665.

58. Sklenar V., Piotto M., Leppik R., Saudek V. *J. Mag. Reson. Series A.* **1993**, 102, 241-245.

59. Liu M., Mao X., Ye C., Huang H., Nicholson J. K., Lindon J. C. Improved WATERGATE pulse sequences for solvent suppression in NMR spectroscopy. *J. Mag. Reson.* **1998**, 132, 125-129.

60. Goddard T. D., Kneller D. G. *Sparky* 3.112 University o California, San Francisco, **2006**.

Bibliography

1. Albericio F. Orthogonal protecting groups for N^{α}-amino and C-terminal carboxy functions in solid-phase peptide synthesis. *Biopolymers* **2000**, 55, 123-139.

2. Altman K. L. Drug-resistant TB in South Africa draws attention from U. N. *New York Times* September 6, **2006**.

3. Aravinda S.., Shamala N., Roy R. S., Balaram P. Non-protein amino acids in peptide design. *Proc. Indian Acad. Sci.* **2003**, 115, 373-400.

4. Arkin M. R., Wells J. A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* **2004**, 3, 301-317.

5. Ashworth J., Havranek J. J., Duarte C. M., Sussman D., Monnat R. J. Jr., Stoddard B. L., Baker D. Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature*, **2006**, 441, 656-659.

6. Baker D. Prediction and design of macromolecular structures and interactions. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2006**, 361, 459-463.

7. Banerjee R., Basu G., Chene P., Roy S. Aib-based peptide backbone as scaffolds for helical peptide mimics. *J. Peptide Res.* **2002**, 60, 88–94.

8. Barder T. E., Walker S. D., Martinelli J. R., Buchwald S. L. Catalysts for Suzuki-Miyaura coupling processes: scope and studies of the effect of ligand structure. *J. Am. Chem. Soc.* **2005**, 127, 4685-4696.

9. Barlow D. J. Peptide drug design. *Smith and Williams' Introduction to the principles of drug design and action (4th Ed.)* **2006**, 327-354.

10. Bastiaens P. I., Majoul I. V., Verveer P. J., Soling H. D., Jovin T. M. Imaging the intracellular trafficking and state of the AB5 quaternary structure of cholera toxin. *EMBO J.* **1996**, 15, 4246-4253.

11. Bax A., Davis D. G. *J. Mag. Reson.* **1985**, 63, 207-213.

12. Bellanda M., Peggion E., Burgi R., van Gunsteren W., Mammi S. Conformational study of an Aib-rich peptide in DMSO by NMR. *J. Peptide Res.* **2001**, 57, 97-106.

13. Berg J. M., Tymoczko J. L., Stryler L., Clarke N. D. *Biochemistry (5^{th} Ed.)* W. H. Freeman and Company, New York, **2002**.

14. Berg T. Modulation of protein-protein interactions with small organic molecules. *Angew. Chem. Int. Ed. Engl.* **2003**, 42, 2462-2481.

15. Blumberg H. M., Leonard M. K. Jr., Jasmer R. M. Update on the treatment of tuberculosis and latent tuberculosis infection. *JAMA* **2005**, 293, 2776-2784.

16. Borman S. Drugs by design. *Chem. Eng. News* **2005**, 83, 28-30.

17. Bosch J., Turley S., Daly T. M., Bogh S. M., Villasmil M. L., Roach C., Zhou N., Morrisey J. M., Vaidya A. B., Bergman L. W., Hol W. G. Structure of the MTIP-MyoA complex, a key component of the malaria parasite invasion motor. *Proc. Natl. Acad. Sci. USA* **2006**, 103, 4852-4857.

18. Boussard C., Klimkait T., Mahmood N., Pritchard M., Gilbert I. H. Design, synthesis and evaluation of potential inhibitors of HIV gp120-CD4 interactions. *Bioorg. Med. Chem. Lett.* **2004**, 14:2673-2676.

19. Bowen W. P., Jerman J. C. Nonlinear regression using spreadsheets. *Trends Pharmacol. Sci.* **1995**, 16, 413-417.

20. Bracken C., Gulyras J., Taylor J. W., Baum J. Synthesis and nuclear magnetic resonance structure determination of an α-helix, bicyclic, lactam-bridged hexapeptide. *J. Am. Chem. Soc.* **1994**, 116, 6431-6432.

21. Brorson S. H. Bovine serum albumin (BSA) as a reagent against non-specific immunogold labeling on LR-White and epoxy resin. *Micron* **1997**, 28, 189-195.

22. Cabezas E., Satterthwait A. C. The hydrogen bond mimic approach: solid-phase synthesis of a peptide stabilized as an α-helix with a hydrazone link. *J. Am. Chem. Soc.* **1999**, 121, 3862-3875.

23. Chakrabartty A., Doig A. J., Baldwin R. L. Helix capping propensities in peptides parallel those in proteins. *Proc. Natl. Acad. Sci. USA* **1993**, 90, 11332-11336.

24. Chemical Computing Group. MOE. Montreal, Quebec, Canada, **2003**.

25. Chen Y. H., Yang T. Y., Chau K. H. Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry*, **1974**, 13, 3350-3359.

26. Chene P. Drugs targeting protein-protein interactions. *ChemMedChem.* **2006**, 4, 400-411.

27. Chene P. Inhibition of the p53-MDM2 interactions: Targeting a protein–protein interface. *Mol. Cancer Res.* **2004**, 2, 20-28.

28. Chene P. p53 as a drug target in cancer therapy. *Expert Opin. Ther. Patents* **2001**, 11, 923-935.

29. Choo Y., Klug A. Designing DNA-binding proteins on the surface of filamentous phage. *Curr. Opin. Biotechnol.* **1995**, 6, 431-436.

30. Cieplak W. Jr., Mead D. J., Messer R. J. and Grant C. C. Site-directed mutagenic alteration of potential active-site residues of the A subunit of Escherichia coli heat-labile enterotoxin. Evidence for a catalytic role for glutamic acid 112. *J. Biol. Chem.* **1995**, 270, 30545-30550.

31. Clackson T., Wells J. A. A hot spot of binding energy in a hormone-receptor interface. *Science* **1995**, 267, 383-386.

32. Clore G. M., Gronenborn A. M. Structures of larger proteins in solution: three- and four-dimensional heteronuclear NMR spectroscopy. *Science* **1991**, 252, 1390-1399.

33. Cochran A. G. Antagonists of protein-protein interactions. *Chem. Biol.* **2000**, 7, 85-94.

34. Cordon-cardo C., Latres E., Drobnjak M., Oliva M. R., Pollack D., Woodruff J. M., Marechal V., Chen J. D., Brennan M. F., Levine A. J. Molecular abnormalities of mdm2 and p53 genes in adult soft tissue sarcomas. *Cancer Res.* **1994**, 54, 794-799.

35. Cosma C. L., Sherman D. R., Ramakrishnan L. The secret lives of the pathogenic mycobacteria. *Annu. Rev. Microbiol.* **2003**, 57, 641-676.

36. Cuthbertson A., Indrevoll B. Regioseletive formation, using orthogonal cysteine protection, of an α-conotoxin dimer peptide containing four disulfide bonds. *Org. lett.* **2003**, 5, 2955-2957.

37. Dall'Acqua W., Carter P. Antibody engineering. *Curr. Opin. Struct. Biol.* **1998**, 8, 443-450.

38. Davis A. M., Teague S. J., Kleywegt G. J. Application and limitations of X-ray crystallographic data in structure-based ligand and drug design. *Angew. Chem. Int. Ed. Engl.* **2003**, 42, 2718-2736.

39. De S. N. Enterotoxicity of bacteria-free culture-filtrate of *Vibrio cholerae*. *Nature* **1959**, 183, 1533-1534.

40. Debouck C. The HIV-1 protease as a therapeutic target for AIDS. *AIDS Res. Hum. Retroviruses.* **1992**, 8, 153-164.

41. Demko Z. P., Sharpless K. B. A click chemistry approach to tetrazoles by Huisgen 1,3-dipolar cycloaddition: synthesis of 5-acyltetrazoles from azides and acyl cyanides. *Angew. Chem. Int. Ed. Engl.* **2002**, 41, 2113-2116.

42. DeVasher R. B., Moore L. R., Shaughnessy K. H. Aqueous-phase, palladium-catalyzed cross-coupling of aryl bromides under mild conditions, using water-soluble, sterically demanding alkylphosphines. *J. Org. Chem.* **2004**, 69, 7919-7927.

43. Dutta N. K., Panse M. W., Kulkarni D. R. Role of cholera a toxin in experimental cholera. *J. Bacteriol.* **1959**, 78, 594-595.

44. Dziejman M., Balon E., Boyd D., Fraser C. M., Heidelberg J. F., Mekalanos J. J. Comparative genomic analysis of Vibrio cholerae: genes that correlate with cholera endemic and pandemic disease. *Proc. Natl. Acad. Sci. USA* **2002**, 99, 1556-1561.

45. Ernst J. T., Becerril J., Park H. S., Yin H., Hamilton A. D. Design and application of an alpha-helix-mimetic scaffold based on an oligoamide-foldamer strategy: antagonism of the Bak BH3/Bcl-xL complex. *Angew. Chem. Int. Ed. Engl.* **2003**, 5, 535-539.

46. Ernst J. T., Kutzki O., Debnath A. K., Jiang S., Lu H., Hamilton A. D. Design of a protein surface antagonist based on alpha-helix mimicry: inhibition of gp41 assembly and viral fusion. *Angew. Chem. Int. Ed. Engl.* **2002**, 41, 278-281.

47. Falcovitz-Gerassi, Y., Tsinberg P. Alleviation of non -specific binding in microarray assays. *U.S. Pat. Appl. Publ.* US 2006040407, **2006**.

48. Fan E., Merritt E. A., Verlinde C. L. M. J., Hol, W. G. J. AB5 toxins: structures and inhibitor design. *Curr. Opin. Struct. Biol.* **2000**, 10, 680-686.

49. Fan E., O'Neal C. J., Mitchell D. D., Robien M. A., Zhang Z., Pickens J. C., Tan X. J., Korotkov K., Roach C., Krumm B., Verlinde C. L., Merritt E. A., Hol W. G. Structural biology and structure-based inhibitor design of cholera toxin and heat-labile enterotoxin. *Int. J. Med. Microbiol.* **2004**, 294, 217-223.

50. Fang H., Kaur G., Yan J., Wang B. An efficient synthesis of sterically hindered arylboronic acids. *Tet. Lett.* **2005**, 46, 1671-1674.

51. Feil I. K., Platas A. A., van den Akker F., Reddy R., Merritt E. A., Storm D. R., Hol W. G. Stepwise transplantation of an active site loop between heat-labile enterotoxins LT-II and LT-I and characterization of the obtained hybrid toxins. *Protein Eng.* **1998**, 11, 1103-1109.

52. Fersht A. A guide to enzyme catalysis and protein folding. W. H. Freeman and Company, New York, **1999**.

53. Finkelstein R. A., Norris H. T., Dutta N. K. Pathogenesis experimental cholera in infant rabbits. *J. Infect. Dis.* **1964**, 114, 203-216.

54. Florczyk M. A., McCue L. A., Purkayastha A., Currenti E., Wolin M. J., McDonough K. A. A family of *acr*-coregulated *Mycobacterium tuberculosis* genes shares a common DNA motif and requires Rv3133c (dosR or devR) for expression. *Infect. Immun.* **2003**, 71, 5332-5343.

55. Fraser M. E., Chernaia M. M., Kozlov Y. V., James M. N. Crystal structure of the holotoxin from Shigella dysenteriae at 2.5 A resolution. *Nat. Struct. Biol.* **1994**, 1, 59-64.

56. Frieden T. R., Sterline T. R., Munsiff S. S., Watt C. J., Dye C. Tuberculosis. *Lancet* **2003**, 362, 887-899.

57. Fujii N., Haresco J. J., Novak K. A. P., Stokoe D., Kuntz I. D., Guy R. K. A selective irreversible inhibitor targeting a PDZ protein interaction domain. *J. Am. Chem. Soc.* **2003**, 125, 12074-12075.

58. Gadek T. R., Nicholas J. B. Small molecule antagonists of proteins. *Biochem. Pharmacol.* **2003**, 65,1-8.

59. Galloway T. S., Van Heyningen S. Binding of NAD+ by cholera toxin. *Biochemistry* **1987**, 244, 225-230.

60. Gellman S. H. Foldamers: A Manifesto. *Acc. Chem. Res.* **1998**, 31, 173-180.

61. Gill D. M. Involvement of nicotinamide adenine dinucleotide in the action of cholera toxin *in vitro*. *Proc. Nal. Acad. Sci. USA* **1975**, 72, 2064-2068.

62. Goddard T. D., Kneller D. G. Sparky 3.112 University o California, San Francisco, **2006**.

63. GOLD Version 1.2, [http://www.ccdc.cam.ac.uk/products/life_sciences/gold/], **2003**.

64. Green L. S., Jellinek D., Bell C., Beebe L. A., Feistner B. D., Gill S. C., Jucker F. M., Janjic N. Nuclease-resistant nucleic acid ligands to vascular permeability factor/vascular endothelial growth factor. *Chem. Biol.* **1995**, 2, 683-695.

65. Greenfield A.A., Butera J. A., Caufield C. E. Convenient synthesis of functionalized terphenyls. *Tet. Lett.* **2003**, 44, 2729-2732.

66. Greenlee W. J. Renin inhibitors. *Pharm. Res.* **1987**, 4, 364-374.

67. Gutierrez-Puente Y., Zapata-Benavides P., Tari A. M., Lopez-Berestein G. Bcl-2-related antisense therapy. *Semin. Oncol.* **2002**, 29, 71-76.

68. Gwaltney S. L. II Design and synthesis of potent, selective, and orally efficacious DPP4 inhibitors accerated by high-throughput structural biology. *30th National Medicinal Chemistry Symposium*, June 25-29, Seattle, **2006**.

69. Halgren T. A., Murphy R. B., Friesner R. A., Beard H. S., Frye L. L., Pollard W. T., Banks J. L. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, 47, 1739–1749.

70. Hamman J. H., Enslin G. M., Kotze A. F. Oral delivery of peptide drugs: barriers and developments. *BioDrugs* **2005**, 19, 165-177.

71. Haupt Y., Maya R., Kazaz A., Oren M. Mdm2 promotes the rapid degradation of p53. *Nature* **1997**, 387, 296-299.

72. Henry C. M. Structure-based drug design. *Chem. Eng. News* **2001**, 79, 69-74.

73. Hipfner D. R., Cohen S. M. Connecting proliferation and apoptosis in development and disease. *Nat. Rev. Mol. Cell Biol.* **2004**, 5, 805-815.

74. Hol W. G. Protein crystallography and computer graphics-toward rational drug design. *Angew. Chem. Int. Ed. Engl.* **1986**, 25, 767-778.

75. Hol W.G. The role of the α-helix dipole in protein structure and function. *Prog. Biophys. Mol. Biol.* **1985**, 45, 149-195.

76. Honda R., Tanaka H., Yasuda H. Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53. *FEBS Lett.* **1997**, 420, 25-27.

77. Horal P., Svennerholm B., Jeansson S., Rymo L., Hall W.W. & Vahlne A. Continuous epitopes of the human immunodeficiency virus type 1 (HIV-1) transmembrane glycoprotein and reactivity of human sera to synthetic peptides representing various HIV-1 isolates. *J. Virol.* **1991**, 65, 2718–2723.

78. Hutchison C. A. III, Phillips S., Edgell M. H., Gillam S., Jahnke P., Smith M. Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.* **1978**, 253, 6551-6560.

79. Hwang T. L., Shaka A. J. Cross relaxation without TOCSY: transverse rotating-frame overhauser effect spectroscopy. *J. Am. Chem. Soc.* **1992**, 114, 3157-3159.

80. Ishiyama T., Murata M., Miyaura N. Palladium (0)-catalyzed cross-coupling reaction of alkoxydiboron with haloarenes: a direct procedure for arylboronic esters. *J. Org. Chem.* **1995**, 60, 7508-7510.

81. Jackson D. Y., King D. S., Chmielewski J., Singh S., Schultz P. G. General approach to the synthesis of short .alpha.-helical peptides. *J. Am. Chem. Soc.* **1991**, 24, 9391-9392.

82. Jennings L. D., Foreman K. W., Rush T. S., Tsao D. H. H., Mosyak L., Li Y. H., Sukhdeo M. N., Ding W. D., Dushin E. G., Kenny C. H. Design and synthesis of indolo[2,3-alpha]quinolizin-7-one inhibitors of the ZipA-FtsZ interaction. *Bioorg. Med. Chem. Lett.* **2004**, 14, 1427-1431.

83. Jennings L. D., Foreman K. W., Rush T. S., Tsao D. H. H., Mosyak L., Kincaid S. L., Sukhdeo M. N., Sutherland A.G., Ding W. D., Kenny C. H. Combinatorial synthesis of substituted 3-(2-indolyl)piperidines and 2-phenyl indoles as inhibitors of Zip A-FtsZ interaction. *Bioorg. Med. Chem.* **2004**, 12, 5115-5131.

84. Ji T., Lee M., Pruitt S. C., Hangauer D. G. Privileged scaffolds for blocking protein-protein interactions: 1,4-disubstituted naphthalene antagonists of transcription factor complex HOX-PBX/DNA. *Bioorg. Med. Chem. Lett.* **2004**, 14, 3875-3879.

85. Jobling M. G., R. K. Holmes. Biological and biochemical characterization of variant A subunits of cholera toxin constructed by site-directed mutagenesis. *J. Bacteriol.* **2001**, 183, 4024-4032.

86. Jones S., Thornton J. M. Searching for functional sites in protein structures. *Proc. Natl. Acad. Sci. USA* **1996**, 93, 13-20.

87. Jordan J. D., Landau E. M., Iyengar R. Signaling networks: the origins of cellular multitasking. *Cell* **2000**, 103, 193-200.

88. Kahn R. A., Fu H., Roy C. R. Cellular hijacking: a common strategy for microbial infection. *Trend in Biochem. Sci.* **2002**, 27, 308-314.

89. Kastantin M. and Tirrell M. Entropic stabilization of alpha-helix formation in peptide amphiphiles. *Poster presented at the MROP conference*, Jan. 31, **2005**.

90. Kearsley S. K., Underwood D. J., Sheridan R. P., Miller M. D. Flexibases: a way to enhance the use of molecular docking methods. *J. Comput. Aided Mol. Des.* **1994**, 8, 565–582.

91. Kemp D. S., Allen T. J., Oslick S. L. The energetics of helix formation by short templated peptides in aqueous solution. 1. Characterization of the reporting helical template Ac-Hel1. *J. Am. Chem. Soc.* **1995**, 117, 6641-6657.

92. Kim I. C., Hamilton A. D. Diphenylindane-Based Proteomimetics Reproduce the Projection of the *i*, *i*+3, *i*+4, and *i*+7 Residues on an α-Helix. *Org. Lett.* **2006**, 8, 1751-1754.

93. Kitchen D. B., Decornez H., Furr J. R., Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* **2004**, 3, 935-949.

94. Klasa R. J., Gillum A. M., Klem R. E., Frankel S. R. Oblimersen bcl-2 antisense: facilitating apoptosis in anticancer treatment. *Antisense Nucleic Acid Drug Dev.* **2002**, 12, 193-213.

95. Kolb H. C., Finn M. G., Sharpless K. B. Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem. Int. Ed. Engl.* **2001**, 40, 2004-2021.

96. Kolb H. C., Sharpless K. B. The growing impact of click chemistry on drug discovery. *Drug Dicovery Today* **2003**, 8, 1128-1137.

97. Kononenko V. L., Shimkus J. K., Giddings J. C., Myers M. N. Feasibility studies on photophoretic effects in field-flow fractionation of particles. *J. Liq. Chromatogr. Relat. Technol.* **1997**, 20, 2907-2929.

98. Kotha S., Lahiri K. Post-assembly peptide modifications by chemical methods. *Curr. Med. Chem.* **2005**, 12, 849-875.

99. Kresheck G. C., Vitello L. B., Erman J. E. Calorimetric studies on the interaction of horse ferricytochrome c and yeast cytochrome c peroxidase. *Biochemistry* **1995**, 34, 8398-8405.

100. Ku J., Schultz P. G. Alternate protein frameworks for molecular recognition. *Proc. Natl. Acad. Sci. USA* **1995**, 92, 6552-6556.

101. Kubbutat M. H. G., Jones S. N., Vousden K. H. Regulation of p53 stability by Mdm2. *Nature* **1997**, 387, 299-303.

102. Kuhn P., Wilson K., Patch M. G., Stevens R. C. The genesis of high-throughput structure-based drug discovery using protein crystallography. *Curr. Opin. Chem. Biol.* **2002**, 6, 704-710.

103. Kuntz I. D. Structure-based strategies for drug design and discovery. *Science* **1992**, 257, 1078-1082.

104. Kuntz I. D., Blaney J. M., Oatley S. J., Langridge R., Ferrin T. E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **1982**, 161, 269-288.

105. Kutzki O., Park H. S., Ernst J. T., Orner B. P., Yin H., Hamilton A. D. Development of a potent Bcl-x(L) antagonist based on alpha-helix mimicry. *J. Am. Chem. Soc.* **2002**, 40, 11838-11839.

106. Lakey J. H., Raggett E. M. Measuring protein–protein interactions. *Curr. Opin. Struct. Biol.* **1998**, 8, 119-123.

107. Larew J. S., Peterson J. E., Graves D. J. Determination of the kinetic mechanism of arginine-specific ADP-ribosyltransferases using a high performance liquid chromatographic assay. *J. Biol. Chem.* **1991**, 266, 52-57.

108. Lencer W. I. Retrograde transport of cholera toxin into the ER of the host cells. *Int. J. Med. Microbiol.* **1994**, 293, 491-494.

109. Lencer W. I., Constable C., Moe S., Jobling M. G., Webb H. M., Ruston S., Madara J. L., Hirst T. R., Holmes R. K. Targeting of cholera toxin and Escherichia coli heat labile toxin in polarized epithelia: role of COOH-terminal KDEL. *J. Cell Biol.* **1995**, 131, 951-962.

110. Li J., Zhang G., Zhang Z., Fan E. TFA-sensitive arylsulfonylthiourea-assisted synthesis of N,N'-substituted guanidines. *J. Org. Chem.* **2003**, 68, 1611-1614.

111. Li P., Roller P. P. Cyclization strategies in peptide derived drug design. *Curr. Top. Med. Chem.* **2002**, 2, 325-341.

112. Li S., Armstrong C. M., Bertin N., Ge H., Milstein S., Boxem M., Vidalain P. O., Han J. D., Chesnau A., Hao T., Goldberg D. S., Li N., Martinez M., Rual J. F., Lamesch P., Xu L., Tewari M., Wong S. L., Zhang L. V., Berriz G. F., Jacotot L., Vaglio P., Reboul J., Hirozane-Kishikawa T., Li Q., Gabel W., Elewa A., Baumgartner B., Rose D. J., Yu H., Bosak S., Sequerra R., Fraser A., Mango S. E., Saxton W. M., Strome S., Van Den Heuvel S., Piano F., Vandenhaute J., Sardet C., Gerstein M., Doucette-Stamm L., Gunsalus K. C., Harper J. W., Cusick E., Roth F. P., Hill D. E., Vidal M. A map of the interactome network of the metazoan C. elegans. *Science* **2004**, 303, 540– 543.

113. Liang N., Mauk A. G., Pielak G. J., Johnson J. A., Smith M., Hoffman B. M. Regulation of interprotein electron transfer by residue 82 of yeast cytochrome c. *Science* **1988**, 240, 311-313.

114. Liu J. Natural terphenyls: developments since 1877. *Chem. Rev.* **2006**, 106, 2209-2223.

115. Liu M., Mao X., Ye C., Huang H., Nicholson J. K., Lindon J. C. Improved WATERGATE pulse sequences for solvent suppression in NMR spectroscopy. *J. Mag. Reson.* **1998**, 132, 125-129.

116. Loesberg C., van Rooij H., Smets L. A. Meta-iodobenzylguanidine (MIBG), a novel high-affinity substrate for cholera toxin that interferes with cellular mono(ADP-ribosylation). *Biochim. Biophys. Acta* **1990**, 1037, 92-99.

117. Luo P., Baldwin L. R. Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* **1997**, 36, 8413-8421.

118. Luo P., Braddock D. T., Subramanian R. M., Meredith S. C., Lynn D. G. Structural and thermal dynamic characterization of a bioactive peptide model of apolipoprotein E: side-chain lactam bridges to constrain the conformation. *Biochemistry* **1994**, 33, 12367-12377.

119. Maeda M., Melnyk R. A., Partridge A. W., Liu L. P., Deber C. M.. Transmembrane segment peptides with double D-amino acid replacements: helicity, hydrophobicity, and antimicrobial activity. *Biopolymers* **2003**, 71, 77-84.

120. Maeyama K., Yonezawa N. Synthesis of substituted m-terphenyls. *Recent Res. Dev. Org. Chem.* **2003**, 7, 63-69.

121. Maeyama K., Yonezawa N. Synthesis of substituted o- terphenyl via electrophilic aromatic substitution and other routes. *Recent Res. Dev. Org. Chem.* **2003**, 7, 53-61.

122. Majoul I. V., Bastiaens P. I., Soling H. D. Transport of an external Lys-Asp-Glu-Leu (KDEL) protein from the plasma membrane to the endoplasmic reticulum: studies with cholera toxin in Vero cells. *J. Cell Biol.* **1996**, 133, 777-789.

123. Makhatadze G. I., Privalov P. L. Energetics of protein structure. *Adv. Protein Chem.* **1995**, 48, 307-425.

124. Malhotra V., Sharma D., Ramanathan V. D., Shakila H., Saini D. K., Chakravorty S., Das T. K., Li Q., Silver R. F., Narayanan P. R., Tyagi J. S. Disruption of response regulator gene, *devR*, leads to attenuation in virulence of *Mycobacterium tuberculosis*. *FEMS Microbiol. Lett.* **2004**, 231, 237-245.

125. Manimala J. C., Anslyn E. V. Solid-phase synthesis of guanidinium derivatives from thiourea functionalities. *Eur. J. Org. Chem.* **2002**, 3909-3922.

126. Martin F., Toniatti C., Salvati A. L., Venturini S., Ciliberto G., Cortese R., Sollazzo M. The affinity-selection of a minibody polypeptide inhibitor of human interleukin-6. *EMBO J.* **1994**, 13, 5303-5309.

127. McMartin C., Bohacek R. S. QXP: Powerful, rapid computer algrithms for structure-based drug design. *J. Comput. Aided Mol. Design* **1997**, 11, 333-344.

128. McMartin C., Bohacek R. S. QXP: Powerful, rapid computer algrithms for structure-based drug design. *J. Comput. Aided Mol. Design* **1997**, 11, 333-344.

129. McNally V. A., Gbaj A., Douglas K. T., Stratford I. J., Jaffar M., Freeman S., Bryce R. A. Identification of a novel class of inhibitor of human and Escherichia coli thymidine phosphorylase by *in silico* screening. *Bio. Med. Chem. Lett.* **2003**, 13, 3705–3709.

130. Mekalanos J. J., Collier R. J., Romig W. R. Enzymic activity of cholera toxin. I. New method of assay and the mechanism of ADP-ribosyl transfer. *J. Biol. Chem.* **1979**, 254, 5849-5854.

131. Meng E. C., Shoichet B. K., Kuntz I. D. Automated docking with grid-based energy evaluation. *J. Comp. Chem.* **1992**, 13, 505-524.

132. Merrit E. A., Hol W. G. AB_5 toxins. *Curr. Opin. Struct. Biol.* **1995**, 5, 165-171.

133. Minoru A. *Spectrochimica Acta, Part A* **1984**, 40, 367-371.

134. Mizejewski G. J. Peptides as receptor ligand drugs and their relationship to G-coupled signal transduction. *Expert Opin. Invest. Drugs* **2001**, 10, 1063-1073.

135. Moss J., Manganiello V. C., Vaughan M. Hydrolysis of nicotinamide adenine dinucleotide by choleragen and its A protomer: possible role in the activation of adenylate cyclase. *Proc. Natl. Acad. Sci. USA* **1976**, 73, 4424-4427.

136. Narayanan J., Hartman P. A., and Graves D. J. Assay of heat-labile enterotoxins by their ADP-ribosyltransferase activities. *J. Clin. Microbiol.* **1989**, 27, 2414-2419.

137. Nathan C., Shiloh M. U. Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proc. Natl. Acad. Sci. USA* **2000**, 97, 8841-8848.

138. Ni W., Fang H., Springsteen G., Wang B. The design of boronic acid spectroscopic reporter compounds by taking advantage of the pK_a-lowering effect of diol binding: nitrophenol-based color reporters for diols. *J. Org. Chem.* **2004**, 69, 1999-2007.

139. Nooren I. M. A., Thornton J. M. Diversity of protein-protein interactions. *EMBO J.* **2003**, 22, 3486-3492.

140. Oliner J. D., Pietenpol J. A., Thiagalingam S., Gvuris J., Kinzler K. W., Vogelstein B. Oncoprotein MDM2 conceals the activation domain of tumour suppressor p53. *Nature* **1993**, 362, 857-860.

141. Olson A. J., Goodsell D. S. Automated docking in crystallography: analysis of the substrates of aconitase. *Proteins* **1993**, 17, 1-10.

142. O'Neal C. J., Amaya E. I., Jobling M. G., Holmes R. K., Hol W. G. Crystal structures of an intrinsically active cholera toxin mutant yield insight into the toxin activation mechanism. *Biochemistry* **2004**, 43, 3772-3782.

143. O'Neal C. J., Jobling M. G., Holmes R. K., Hol W. G. Structural basis for the activation of cholera toxin by human ARF6-GTP. *Science* **2005**, 309, 1093-1096.

144. Oneyama C., Agatsuma T., Kanda Y., Nakano H., Sharma S. V., Nakano S., Narazaki F., Tatsuta K. Synthetic inhibitors of proline-rich ligand-mediated protein-protein interaction: potent analogs of UCS15A. *Chem. Biol.* **2003**, 10, 443-451.

145. Orner B. P., Ernst J. T., Hamilton A. D. Toward proteomimetics: terphenyl derivatives as structural and functional mimics of extended regions of an alpha-helix. *J. Am. Chem. Soc.* **2001**, 123, 5382-5383.

146. Orner B. P., Hamilton, A. D. The guanidinium group in molecular recognition: design and synthetic approaches. *J. Inclusion Phenom. Macrocyclic Chem.* **2001**, 41, 141-147.

147. Osborne J. C. Jr., Stanley S. J., Moss J. Kinetic mechanisms of two NAD: arginine ADP-ribosyltransferases: the soluable, salt-stimulated transferase from turkey erythrocytes and chleragen, a toxin from *Vibrio cholerae*. *Biochemistry* **1985**, 5235-5240.

148. Pagliaro L., Felding J., Audouze K., Nielsen S. J., Terry R. B., Krog-Jensen C., Butcher S. Emerging classes of protein-protein interaction inhibitors and new tools for their development. *Curr. Opin. Chem. Biol.* **2004**, 8, 442-449.

149. Papavassiliou A. G., Bohmann D. Optimization of the signal-to-noise ratio in south- western assays by using lipid-free BSA as blocking reagent. Nucleic Acids Res. **1992**, 20, 4365-4366.

150. Parish T., Smith D. A., Kendall S., Casali N., Bancroft G. J., Stoker N. G. Deletion of two-component regulatory systems increases the virulence of *Mycobacterium tuberculosis*. *Infect. Immun.* **2003**, 71, 1134-1140.

151. Park H. D., Guinn K. M., Harrell M. I., Liao R., Voskuil M. I., Tompa M., Schoolnik G. K., Sherman, D. R. Rv3133c/dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **2003**, 48, 833-843.

152. Parrish N. M., Dick J. D., Bishai W. R. Mechanisms of latency in *Mycobacterium tuberculosis*. *Trends Microbiol.* **1998**, 6, 107-112.

153. Pellecchia M., Sem D. S., Wuthrich K. NMR in drug discovery. *Nat. Rev.* **2002**, 211-219.

154. Pellegrini M., Haynor D., Johnson J. M. Protein interaction networks. *Expert Rev. Proteomics* **2004**, 1, 239-249.

155. Peterfi Z., Kocsis B. Comparison of blocking agents for an ELISA for LPS. *J. Immunoassay* **2000**, 21, 341-354.

156. Piotto M., Saudek V., Sklenár V. A simple modification to the WATERGATE water suppression scheme. *J. Biomol. NMR* **1992**, 2, 661–665.

157. Pugsley A. P. The complete general secretory pathway in gram-negative bacteria. *Microbiol. Rev.* **1993**, 57, 50-108.

158. Pulungsih S. P., Punjabi N. H., Rafli K., Rifajati A., Kumala S., Simanjuntak C. H. Lesmana M., Subekti D., Sutoto F. O. Standard WHO-ORS versus reduced-osmolarity ORS in the management of cholera patients. *J. Health Popul. Nutr.* **2006**, 24, 107-112.

159. Rarey M., Kramer B., Lengauer T., Klebe G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, 261, 470-489.

160. Ringia E. A. T., Schramm V. L. Transition states and inhibitors of the purine nucleoside phosphorylase family. *Curr. Topics Med. Chem.* **2005**, 5, 1237-1258.

161. Rising K. A., Schramm V. L. Transition state analysis of NAD^+ hydrolysis by the cholera toxin catalytic subunit. *J. Am. Chem. Soc.* **1997**, 119, 27-37.

162. Roberts D. M., Liao R. P., Wisedchaisri G., Hol W. G., Sherman D. R. Two sensor kinases contribute to the hypoxic response of *Mycobacterium tuberculosis*. *J. Biol. Chem.* **2004**, 279, 23082-23087.

163. Robien M. A., Krumm B. E., Sandkvist M., Hol W. G. Crystal structure of the extracellular protein secretion NTPase EpsE of Vibrio cholerae. *J. Mol. Biol.* **2003**, 333, 657-674.

164. Robinshaw J. D., Russel D. W., Harris B. A., Smigel M. D., Gilman A. G. Deduced primary structure of the alpha subunit of the GTP-binding stimulatory protein of adenylate cyclase. *Proc. Natl. Acad. Sci. USA* **1986**, 83, 1251-1255.

165. Rodriguez A. L., Tamrazi A., Collins M. L., Katzenellenbogen J. A. Design, synthesis, and in vitro biological evaluation of small molecule inhibitors of estrogen receptor a coactivator binding. *J. Med. Chem.* **2004**, 47, 600-611.

166. Sack D. A., Sack R. B., Chaignat C. L. Getting serious about cholera. *N. Engl. J. Med.* **2006**, 355, 649-651.

167. Saini D. K., Malhotra V., Dey D., Pant N., Das T. K., Tyagi J. S. DevR-DevS is a bona fide two-component system of *Mycobacterium tuberculosis* that is hypoxia-responsive in the absence of the DNA-binding domain of DevR. *Microbiology* **2004**, 150, 865-875.

168. Saini D. K., Malhotra V., Tyagi J. S. Cross talk between DevS sensor kinase homologue, Rv2027c, and DevR response regulator of *Mycobacterium tuberculosis*. *FEBS Lett.* **2004**, 565, 75-80.

169. Salmond R. J., Luross J. A., Williams N. A. Immune modulation by the cholera-like enterotoxins. *Expert Rev. Mol. Med.* **2002**, 1 October, 1-16.

170. Sarabia F., Chammaa S., Ruiz A. S., Ortiz L. M., Herrera F. J. L. Chemistry and biology of cyclic peptides of medicinal and biological interest. *Curr. Med. Chem.* **2004**, 11, 1309-1332.

171. Schmiedeberg N., Kessler H. Reversible backbone protection enables combinatorial solid-phase ring-closing metathesis reaction (RCM) in peptides. *Org. Lett.* **2002**, 4, 59-62.

172. Schmitz A., Herrgen H., Winkeler A., Herzog V. Cholera toxin is exported from microsomes by the Sec61p complex. *J. Cell Biol.* **2000**, 148, 1203-1212.

173. Schramm V. L. Enzymatic Transition State Poise and Transition State Analogues. *Acc. Chem. Res.* **2003**, 36, 588-596.

174. Schramm V. L. Enzymatic transition states and transition state analogues. *Curr. Opin. Struct. Biol.* **2005**, 15, 604-613.

175. Schramm V. L. Enzymatic transition states: thermodynamics, dynamics and analog design. *Arch. Biochem. Biophys.* **2005**, 433, 13-26.

176. Schramm V. L., Shi W. Atomic motion in enzymatic reaction coordinates. *Curr. Opin. Struct. Boil.* **2001**, 11, 657-665.

177. Schueler-Furman O., Wang C., Bradley, P., Misura K., Baker D. Progress in modeling of protein structures and interactions. *Science* **2005**, 310, 638-642.

178. Scolnik Y., Portnaya I., Cogan U., Tal S., Haimovitz R., Fridkin M., Elitzur A. C., Deamer D. W., Shinitzky M. Subtle differences in structural transitions between poly-L- and poly-D-amino acids of equal length in water. *Phys. Chem. Chem. Phys.* **2006**, 8, 333-339.

179. Seidler J., McGovern S. L., Doman T. N., Shoichet B. K. Identification and prediction of promiscuous aggregating inhibitors among known drugs. *J. Med. Chem.* **2003**, 46, 4477-4486.

180. Shepherd N., Hoang H. N., Abbenante G., Fairlie D. P. Single turn peptide alpha helices with exceptional stability in water. *J. Am. Chem. Soc.* **2005**, 127, 2974-2983.

181. Sherman D. R., Voskuil M., Schnappinger D., Liao, R. Harrell, M. I., Schoolnik G. K. Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding Alpha-crystallin. *Proc. Natl. Acad. Sci. USA* **2001**, 98, 7534-7539.

182. Shoichet B. K., Kuntz I. D., Bodian D. L. Molecular docking using shape descriptors. *J. Comp. Chem.* **1992**, 13, 380-397.

183. Singh R. Viciu M. S., Kramareva N., Navarro O., Nolan S. P. Simple (imidazol-2-ylidene)-Pd-acetate complexes as effective prcatalysts fpr sterically hindered Susuki-Miyaura couplings. *Org. Lett.* **2005**, 7, 1829-1832.

184. Sixma T. K., Pronk S. E., Kalk K. H., Wartna E. S., van Zanten B. A. M., Witholt B. , Hol, W. G. J. Crystal structure of a cholera toxin-related heat-labile enterotoxin from *E. coli. Nature* **1991**, 351, 371-377.

185. Sklenar V., Piotto M., Leppik R., Saudek V. *J. Mag. Reson. Series A.* **1993**, 102, 241-245.

186. Smith I. *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. *Clin. Microbiol. Rev.* **2003**, 16, 463-496.

187. Smith M. In vitro mutagenesis. *Annu. Rev. of Gen.* **1985**, 19, 423-462.

188. Smith M. Site-directed mutagenesis. *Trends in Biochem. Sci.* **1982**, 7, 440-442.

189. Smith M. The role of chemical synthesis in establishing gene function by in vitro mutagenesis. *Proceedings of the Robert A. Welch Foundation Conference on Chemical Research.* **1985**, 29, 439-455.

190. Smith M., Leung D. W., Gillam S., Astell C. R., Montgomery D. L., Hall, B. D. Sequence of the gene for iso-1-cytochrome c in Saccharomyces cerevisiae. *Cell* **1979**, 16, 753-761.

191. Smyth T. P. Substrate variants versus transition state analogues as noncovalent, reversible enzyme inhibitors. *Bioorg. Med. Chem.* **2004**, 12, 4081-4088.

192. Soman G., Narayanan J., Martin B. L., Graves D. J. Use of substituted (benzylidineamino)guanidines in the study of guanidine group specific ADP-ribosyltrasferase. *Biochemistry* **1986**, 25, 4113-4119.

193. Stein P. E., Boodhoo A., Tyrrell G. J., Brunton J. L., Read R. J. Crystal structure of the cell-binding B oligomer of verotoxin-1 from *E. coli. Nature* **1992**, 355, 748-750.

194. Stewart G. R., Robertson B. D., Young D. B. Tuberculosis: a problem with persistence. *Nat. Rev. Microbiol.* **2003**, 1, 97-105.

195. Sun Y., Ewing T. J. A., Skillman A. G., Kuntz I. D. CombiDOCK: structure-based combinatorial docking and library design. *J. Comput.-Aided Mol. Des.* 1998, 12, 594

196. Tan C., Wei L. H., Ottensmeyer F. P., Goldfine I., Maddux B. A., Yip C. C., Batey R. A., Kotra L. P. Structure-based de novo design of ligands using a three-

dimensional model of the insulin receptor. *Bioorg. Med. Chem. Lett.* **2004**, 14, 1407-1410.

197. Taylor J. W. The synthesis and study of side-chain lactam-bridged peptides. *Biopolymers.* **2002**, 1, 49-75.

198. Taylor J. W. The synthesis and study of side-chain lactam-bridged peptides. *Biopolymers* **2002**, 1, 49-75.

199. Taylor J. W., Jin Q. K., Sbacchi M., Wang L., Belfiore P., Garnier M., Kazantzis A., Kapurniotu A., Zaratin P. F., Scheideler M. A. Side-chain lactam-bridge conformational constraints differentiate the activities of salmon and human calcitonins and reveal a new design concept for potent calcitonin analogues. *J. Med. Chem.* **2002**, 5, 1108-1121.

200. Taylor R. D., Jewsbury P. J., Essex J. W. A review of protein-small molecule docking methods. *J. Comput. Aid. Mol. Design* **2002**, 16, 151–166.

201. Teng Z., Moore J. L., Craig A. R., Hickey G., Schelp C., Wei T. Reduction of non -specific binding in assays. *U.S. Pat. Appl. Publ.* US 2006105472, **2006**.

202. Teter K., Holmes R. K., Schmitz A., Herrgen H., Winkeler A. and Herzog V. Inhibition of endoplasmic reticulum-associated degradation in CHO cells resistant to cholera toxin, Pseudomonas aeruginosa exotoxin A, and ricin. Cholera toxin is exported from microsomes by the Sec61p complex. *Infect. Immun.* **2002**, 70, 6172-6179.

203. ThistleSoft, Inc. FLO96 ThistleSoft, Colebrook, CT, **1996**.

204. Tian Y., Ramesh C. V., Ma X., Naqvi S., Patel T., Cenizal T., Tiscione M., Diaz K., Crea T., Arnold E., Arnold G. F., Taylor J. W. Structure-affinity relationships in the gp41 ELDKWA epitope for the HIV-1 neutralizing monoclonal antibody 2F5: effects of side-chain and backbone modifications and conformational constraints. *J. Pept. Res.* **2002**, 6, 264-76.

205. Toogood P. L. Inhibition of protein-protein association by small molecules: approaches and progress. *J. Med. Chem.* **2002**, 45, 1543-1558.

206. Tsai B., Rapoport T. A. Unfolded cholera toxin is transferred to the ER membrane and released from protein disulfide isomerase upon oxidation by Ero1. *J. Cell Biol.* **2002**, 159, 207-216.

207. Tsai B., Rodighiero C., Lencer W. I., Rapoport T. A. Protein disulfide isomerase acts as a redox-dependent chaperone to unfold cholera toxin. *Cell* **2001**, 104, 937-948.

208. Uchida K., Segawa T., Takahashi Y. Absorption bands in the absorption edge region of p-terphenyl crystals. *J. Lumin.* **1994**, 60-61, 795-797.

209. Ulrich R. Dock around the clock - current status of small molecule docking and scoring. *QSAR & Comb. Sci.* **2006**, 25, 605-615.

210. van den Akker F., Merritt E. A., Pizza M., Domenighini M., Rappuoli R., Hol W. G. The Arg7Lys mutant of heat-labile enterotoxin exhibits great flexibility of active site loop 47-56 of the A subunit. *Biochemistry* **1995**, 34, 10996-11004.

211. Vaughan T. J., Osbourn J. K., Tempest P. R. Human antibodies by design. *Nat. Biotechnol.* **1998**, 16, 535-539.

212. Wallimann P., Kennedy R. J., Kemp D. S. Large circular dichroism ellipticities for N-templated helical polypeptides are inconsistent with currently accepted helicity algorithms. *Angew. Chem. Int. Ed. Engl.* **1999**, 38, 1290-1292.

213. Wayne L. G., Sohaskey C. D. Nonreplicating persistence of *Mycobacterium tuberculosis*. *Annu. Rev. Microbiol.* **2001**, 55, 139-163.

214. Weber G. Thermodynamics of the association and the pressure dissociation of oligomeric proteins. *J. Phys. Chem.* **1993**, 97, 7108-7115.

215. Whittle P. J., Blundell T. L. Protein Structure-based drug design. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, 23, 349-375.

216. Widdus R. Public-private partnerships for health: their main targets, their diversity, and their future directions. *Bulletin of the World Health Organization.* **2001**, 79, 713-720.

217. Williams D. H., Calderone C. T., O_Brien D. P., Zerella R. Changes in motion vs. bonding in positively vs. negatively cooperative interactions. *Chem. Commun.* **2002**, 1266-1267.

218. Williams D. H., Maguire A. J., Tsuzuki W., Westwell M. S. An analysis of the origins of a cooperative binding energy of dimerization. *Science* **1998**, 280, 711-714.

219. Williams D. H., Stephens E., Zhou M. How can enzymes be so efficient? *Chem. Commun.* **2003**, 16, 1973-1976.

220. Williams D. H., Stephens E., Zhou M. Ligand binding energy and catalytic efficiency from improved packing within receptors and enzymes. *J. Mol. Biol.* **2003**, 329, 389-399.

221. Williams S. P., Kuyper L. F., Pearce K. H. Recent applications of protein crystallography and strcture-guided drug design. *Curr. Opin. Chem. Biol.* **2005**, 9, 371-380.

222. Winter G., Fersht A. R., Wilkinson A. J., Zoller M., Smith M. Redesigning enzyme structure by site-directed mutagenesis: tyrosyl tRNA synthetase and ATP binding. *Nature* **1982**, 299, 756-758.

223. Wisedchaisri G., Wu M., Rice A. E., Roberts D. M., Sherman D. R., Hol W. G. Structures of Mycobacterium tuberculosis DosR and DosR-DNA complex involved in gene activation during adaptation to hypoxic latency. *J. Mol. Biol.* **2005**, 354, 630-641.

224. Wishart D. NMR spectroscopy and protein structure determination: Applications to drug discovery and development. *Curr. Pharm. Biotech.* **2005**, 6, 105-120.

225. Wishart D. S., Sykes B. D., Richards F. M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* **1991**, 222, 311-333.

226. Wishart D. S., Sykes B. D., Richards F. M. Simple techniques for the quantification of protein secondary structure by 1H NMR spectroscopy. *FEBS Lett.* **1991**, 293, 72-80.

227. Wishart D. S., Sykes B. D., Richards F. M. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **1992**, 31, 1647-1651.

228. Wolfenden R., Radzicka A. Transition-state analogues. *Curr. Opin. Struct. Biol.* **1991**, 1, 780-787.

229. World Health Organization. Cholera Cases and Deaths Reported to WHO, 2002, [http://www.who.int/emc/diseases/cholera/choltbld2002.html]. *WHO Weekly* **2002**.

230. World Health Organization. Global epidemics and impact of cholera. [http://www.who.int/topics/cholera/impact/en/], **2006**.

231. World Health Organization. Concern over extensive drug resistant TB strains that are virtually untreatable. *News release by WHO.* 5 September, **2006**.

232. World Health Organization. Global tuberculosis control: surveillance, planning, financing. *WHO report 2005*, World Health Organization, Geneva, **2005**.

233. World Health Organization. Intensified action needed to halt a quarter of a million TB/HIV Deaths a year to people living with AIDS. *Joint News Release WHO/International AIDS Society WHO/41.* 16 August, **2006**.

234. Wuthrich K. NMR of proteins and nucleic acids. John Wiley and Sons Inc., USA, **1986**.

235. Wuthrich K. Protein structure determination in solution by nuclear magnetic resonance spectroscopy. *Science* **1989**, 243, 45-50.

236. Yang J., Mou J., Shao Z. Structure and stability of pertussis toxin studied by in situ atomic force microscopy. *FEBS Lett.* **1994**, 338, 89-92.

237. Yin H., Hamilton A. D. Strategies for targeting protein-protein interactions with synthetic agents. *Angew. Chem. Int. Ed. Engl.* **2005**, 27, 4130-4163.

238. Yin H., Hamilton A. D. Terephthalamide derivatives as mimetics of the helical region of Bak peptide target Bcl-xL protein. *Bioorg. Med. Chem. Lett.* **2004**, 6, 1375-1379.

239. Yin H., Lee G. I., Park H. S., Payne G. A., Rodriguez J. M., Sebti S. M., Hamilton A. D. Terphenyl-based helical mimetics that disrupt the p53/HDM2 interaction. *Angew. Chem. Int. Ed. Engl.* **2005**, 18, 2704-2707.

240. Yin H., Lee G. I., Sedey K. A., Rodriguez J. M., Wang H. G., Sebti S. M., Hamilton A. D. Terephthalamide derivatives as mimetics of helical peptides: disruption of the Bcl-x(L)/Bak interaction. *J. Am. Chem. Soc.* **2005**, 15, 5463-5468.

241. Zhang R. G., Scott D. L., Westbrook M. L., Nance S., Spangler B. D., Shipley G. G., Westbrook E. M., *J. Mol. Biol.* **1995**, 251, 563-573.

242. Zhao L., Chmielewski J. Inhibiting protein-protein interactions using designed molecules. *Curr. Opin. Struct. Biol.* **2005**, 1, 31-34.

243. Zhou G., Parikn S. L., Tyler P. C., Evans G. B., Furneaux R. H., Zubkova O. V., Benjes P. A., Schramm V. L. Inhibitors of ADP-ribosylating bacterial toxins based on oxacarbenium ion character at their transitions states. *J. Am. Chem. Soc.* **2004**, 126, 5690-5698.

JOC*Note*

APPENDIX A

TFA-Sensitive Arylsulfonylthiourea-Assisted Synthesis of N,N′-Substituted Guanidines

Jizhen Li, Guangtao Zhang,† Zhongsheng Zhang, and Erkang Fan*

Biomolecular Structure Center, Departments of Biochemistry and Biological Structure, and Department of Chemistry, Box 357742, University of Washington, Seattle, Washington 98195

erkang@u.washington.edu

Received December 4, 2002

Abstract: An efficient synthesis of N,N′-substituted guanidine derivatives was developed via an aromatic sulfonyl-activated thiourea intermediate. The use of certain aromatic sulfonamides, such as PbfNH$_2$, as the key reagent to incorporate a TFA-labile guanidine protection group greatly facilitates solid-phase synthesis of N,N′-substituted guanidine compounds.

The guanidine group plays important roles in biological systems and is therefore a component of many therapeutic agents that aim to mimic or block the function of a guanidine-containing biomolecule.[1] The synthesis of guanidine derivatives has also attracted continued research interests in recent years, resulting in many new efficient synthetic methods and guanidinylation reagents for different classes of guanidine compounds.[2−43]

We are interested in the preparation of N,N′-substituted guanidines using common solid-phase supports that can be conveniently cleaved with TFA treatment. Currently, preparation of this class of guanidine compounds is possible through two major synthetic routes. One route is based on di-Boc (or equivalent TFA-labile group)-activated pyrazole- or triazolecarboxamidines[8,19,20,23,44] or triflyl guanidines,[11,27,45] either in solution or on solid support. In these cases, it usually requires reaction with a primary alcohol under Mitsunobu conditions before the guanidinylation step to produce the desired substitution. However, this may limit the diversity of compounds generated because Mitsunobu conditions work best with primary alcohols. The other widely used route is based on modification of a thiourea, of which the solid-phase methods have been reviewed recently.[2] Common ways of improving reaction efficiency through this type of reaction are the activation of thiourea with electron-withdrawing groups such as alkoxycarbonyl, acyl, triazene, or *N*-aryl.[4,14,15,21,22,26,41] Subsequent guanidinylation is usually done with the assistance of a carbodiimide,[41] heavy metal salts,[7,40] or Mukaiyama reagent.[46] The characteristic of

† Department of Chemistry.

(1) *Guanidino Compounds in Biology and Medicine: 2*; De Deyn, P., Mareseau, B., Quneshi, I. A., Mori, A., Eds.; John Libbey & Co. Limited: London, 1997.
(2) Manimala, J. C.; Anslyn, E. V. *Eur. J. Org. Chem.* **2002**, *2002*, 3909−3922.
(3) Wu, Y.-Q.; Hamilton, S. K.; Wilkinson, D. E.; Hamilton, G. S. *J. Org. Chem.* **2002**, *67*, 7553−7556.
(4) Manimala, J. C.; Anslyn, E. V. *Tetrahedron Lett.* **2002**, *43*, 565−567.
(5) Kilburn, J. P.; Lau, J.; Jones, R. C. F. *Tetrahedron* **2002**, *58*, 1739−1743.
(6) Hopkins, T. P.; Dener, J. M.; Boldi, A. M. *J. Comb. Chem.* **2002**, *4*, 167−174.
(7) Cunha, S.; de Lima, B. R.; de Souza, A. R. *Tetrahedron Lett.* **2002**, *43*, 49−52.
(8) Musiol, H. J.; Moroder, L. *Org. Lett.* **2001**, *3*, 3859−3861.
(9) Mamat, A.; Madalengoitia, J. S. *Org. Lett.* **2001**, *3*, 561−564.
(10) Wermann, K.; Walther, M.; Gunther, W.; Gorls, H.; Anders, E. *J. Org. Chem.* **2001**, *66*, 720−726.
(11) Zapf, C. W.; Creighton, C. J.; Tomioka, M.; Goodman, M. *Org. Lett.* **2001**, *3*, 1133−1136.
(12) Zhang, Y. D.; Kennan, A. J. *Org. Lett.* **2001**, *3*, 2341−2344.
(13) Overman, L. E.; Wolfe, J. P. *J. Org. Chem.* **2001**, *66*, 3167−3175.
(14) Li, M.; Wilson, L. J.; Portlock, D. E. *Tetrahedron Lett.* **2001**, *42*, 2273−2275.
(15) Ghosh, A. K.; Hol, W. G. J.; Fan, E. *J. Org. Chem.* **2001**, *66*, 2161−2164.
(16) Acharya, A. N.; Ostresh, J. M.; Houghten, R. A. *J. Comb. Chem.* **2001**, *3*, 578−589.
(17) del Fresno, M.; El-Faham, A.; Carpino, L. A.; Royo, M.; Albericio, F. *Org. Lett.* **2000**, *2*, 3539−3542.
(18) Katritzky, A. R.; Rogovoy, B. V.; Chassaing, C.; Vvedensky, V. *J. Org. Chem.* **2000**, *65*, 8080−8082.
(19) Chen, J.; Pattarawarapan, M.; Zhang, A. J.; Burgess, K. *J. Comb. Chem.* **2000**, *2*, 276−281.
(20) Pátek, M.; Smrčina, M.; Nakanishi, E.; Izawa, H. *J. Comb. Chem.* **2000**, *2*, 370−377.
(21) Linton, B. R.; Carr, A. J.; Orner, B. P.; Hamilton, A. D. *J. Org. Chem.* **2000**, *65*, 1566−1568.
(22) Dahmen, S.; Bräse, S. *Org. Lett.* **2000**, *2*, 3563−3565.
(23) Yong, Y. F.; Kowalski, J. A.; Thoen, J. C.; Lipton, M. A. *Tetrahedron Lett.* **1999**, *40*, 53−56.
(24) Wilson, L. J.; Klopfenstein, S. R.; Li, M. *Tetrahedron Lett.* **1999**, *40*, 3999−4002.
(25) Chang, J.; Oyelaran, O.; Esser, C. K.; Kath, G. S.; King, G. W.; Uhrig, B. G.; Konteatis, Z.; Kim, R. M.; Chapman, K. T. *Tetrahedron Lett.* **1999**, *40*, 4477−4480.
(26) Barawkar, D. A.; Linkletter, B.; Bruice, T. C. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1517−1520.
(27) Feichtinger, K.; Sings, H. L.; Baker, T. J.; Matthews, K.; Goodman, M. *J. Org. Chem.* **1998**, *63*, 8432−8439.
(28) Josey, J. A.; Tarlton, C. A.; Payne, C. E. *Tetrahedron Lett.* **1998**, *39*, 5899−5902.
(29) Kearney, P. C.; Fernandez, M.; Flygare, J. A. *Tetrahedron Lett.* **1998**, *39*, 2663−2666.
(30) Dodd, D. S.; Wallace, O. B. *Tetrahedron Lett.* **1998**, *39*, 5701−5704.
(31) Valdyanathan, G.; Zalutsky, M. R. *J. Org. Chem.* **1997**, *62*, 4867−4869.
(32) Schlama, T.; Gouverneur, V.; Valleix, A.; Greiner, A.; Toupet, L.; Mioskowski, C. *J. Org. Chem.* **1997**, *62*, 4200−4202.
(33) Robinson, S.; Roskamp, E. J. *Tetrahedron* **1997**, *53*, 6697−6705.
(34) Drewry, D. H.; Gerritz, S. W.; Linn, J. A. *Tetrahedron Lett.* **1997**, *38*, 3377−3380.
(35) Barvian, M. R.; Showalter, H. D. H.; Doherty, A. M. *Tetrahedron Lett.* **1997**, *38*, 6799−6802.
(36) Kent, D. R.; Cody, W. L.; Doherty, A. M. *Tetrahedron Lett.* **1996**, *37*, 8711−8714.
(37) Ramadas, K.; Srinivasan, N. *Tetrahedron Lett.* **1995**, *36*, 2841−2844.
(38) Dodd, D. S.; Kozlowski, A. P. *Tetrahedron Lett.* **1994**, *35*, 977−980.
(39) Burgess, K.; Lim, D.; Ho, K. K.; Ke, C. Y. *J. Org. Chem.* **1994**, *59*, 2179−2185.
(40) Kim, K. S.; Qian, L. G. *Tetrahedron Lett.* **1993**, *34*, 7677−7680.
(41) Poss, M. A.; Iwanowicz, E.; Reid, J. A.; Lin, J.; Gu, Z. X. *Tetrahedron Lett.* **1992**, *33*, 5933−5936.
(42) Maryanoff, C. A.; Stanzione, R. C.; Plampin, J. N.; Mills, J. E. *J. Org. Chem.* **1986**, *51*, 1882−1884.
(43) Oh, K. C.; Lee, H.; Kim, K. T. *Tetrahedron Lett.* **1992**, *33*, 4963−4966.
(44) Bernatowicz, M. S.; Wu, Y. L.; Matsueda, G. R. *Tetrahedron Lett.* **1993**, *34*, 3389−3392.
(45) Feichtinger, K.; Zapf, C.; Sings, H. L.; Goodman, M. *J. Org. Chem.* **1998**, *63*, 3804−3805.
(46) Shibanuma, T.; Shiono, M.; Mukaiyama, T. *Chem. Lett.* **1977**, 575.

10.1021/jo026807m CCC: $25.00
Published on Web 01/22/2003

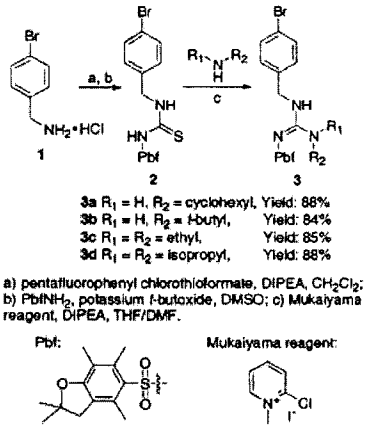
this route is that N,N′-substitution can be introduced with two amine starting materials, although di-Boc-activated isothiourea still requires Mitsunobu conditions to introduce desired substitution patterns.[30] Despite the extensive knowledge about solid-phase synthesis of guanidines, current solid-phase methods for the preparation of N,N′-substituted guanidines starting from two amines only demonstrate limited success.[19,24,25,28,29] Here, we report an extension of current methodologies by exploration of thiourea activation using a single TFA-sensitive aromatic sulfonyl group, along with a comparison to existing methods.

The major goal of our investigation is to achieve in solid-phase synthesis concomitant cleavage of product and removal of the activating arylsulfonyl group, without any loss in overall synthetic efficiency. The question then comes down to (i) if an arylsulfonyl group can sufficiently activate thiourea for guanidine synthesis and (ii) if a TFA-labile protection group can be generated after the guanidinylating step. In addition, we would like to learn if guanidinylation with arylsulfonylthiourea can be carried out under mild experimental conditions for amines with steric hindrance without the need of using toxic heavy metal salts or excessive heating. We should point out that the recently reported use of bismuth nitrate for promoting guanidinylation[7] should be much more environmentally friendly as compared to the use of $HgCl_2$ or analogous salts; however, such a reaction protocol still requires heating (>18 h) or prolonged time at room temperature (up to 117 h) to achieve satisfactory yields.

Although sulfonylguanidines are known, especially in the patented preparation of herbicides, most of the syntheses were done by preformation of the guanidine moiety followed by sulfonylation with sulfonyl chlorides. A survey of literature reporting use of arylsulfonyl activated thiourea for guanidine synthesis showed, to the best of our knowledge, only three reports in solution synthesis.[47-49] Two of these reports used mercury salts to promote the reaction,[47,48] while the third report showed only one example of arylsulfonylthiourea-based guanidinylation using a carbodiimide.[49] In all cases, the formed arylsulfonyl guanidines were not adaptable to solid-phase synthesis using TFA for cleavage/deprotection. We envisioned that if TFA-sensitive arylsulfonyl groups could be introduced into the thiourea-based synthetic method, then adaptation to resin-supported synthesis should be straightforward. The use of the TFA-sensitive sulfonyl group in guanidine synthesis through other types of guanidinylating reagents has also been reported.[36,39] Since many arylsulfonyl protection groups for the arginine side chain are known in Fmoc chemistry, we demonstrate our synthetic method based on one of the best TFA-sensitive arginine protecting groups: 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf).[50]

Since the original report of carbodiimide-promoted arylsulfonylthiourea-assisted guanidine synthesis did not provide enough information about the reactivity for different amines and only sulfamoyl-based systems were

SCHEME 1

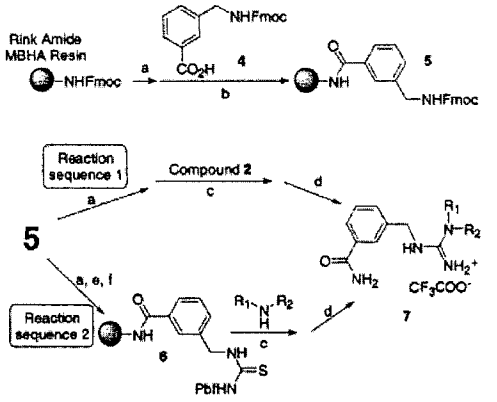


studied in more detail,[49] we first tested the Pbf-thiourea-assisted guanidine formation reaction in solution in the hope of obtaining basic knowledge about the guanidinylating step with a variety of amine nucleophiles. As shown in Scheme 1, a primary amine **1** was first turned into the corresponding pentafluorophenyl thiocarbamate. This allowed for the smooth synthesis of the arylsulfonyl-activated thiourea **2**, using PbfNHK (formed by treating $PbfNH_2$ with potassium *tert*-butoxide) as nucleophile.[48] This preparation of arylsulfonyl modified thiourea has some advantages over an alternative procedure through the use of arylsulfonyl isothiocyanate[48] in that both the starting materials (thiocarbamate and $PbfNH_2$) are fairly stable and easy to handle, while the alternative synthesis of arylsulfonyl-activated thioureas involves the preparation and handling of highly labile arylsulfonyl isothiocyanate, which can react with alcohol without activation under neutral conditions.[51] The resulting compound **2** is an excellent guanidinylating reagent. Treatment of **2** with an amine in the presence of Mukaiyama reagent produced the subsequent guanidine derivatives **3a–d** in very good yields at room temperature in 12–18 h, while most of the product was formed within 1–2 h as judged by TLC. The reaction works for either primary or secondary amine nucleophiles, including the sterically hindered *tert*-butylamine, as well as diethylamine and diisopropylamine, both of which are known to cause problems in other guanidine syntheses performed in solution.[11] In addition, the high efficiency of Mukaiyama reagent in promoting the guanidinylating reaction is also in contrast to its role in sulfamoylthiourea based systems.[49] The transformation was similarly good with the use of EDC in place of Mukaiyama reagent. The use of toxic heavy metal salts or excessive heating proved to be unnecessary.

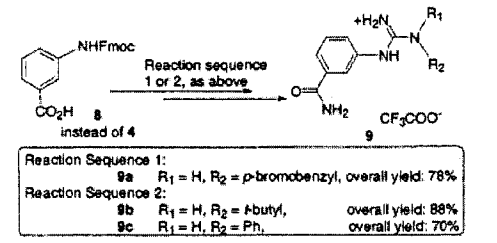
The success in solution transformation led us to investigate a facile solid-phase route for guanidine synthesis via Pbf-activated thiourea, so that free N,N′-substituted guanidines could be obtained directly after TFA cleavage. The adaptation of solution synthesis to

(47) Deprez, P.; Vevert, J.-P. *Synth. Commun.* **1996**, *26*, 4299–4310.
(48) Levallet, C.; Lerpiniere, J.; Ko, S. Y. *Tetrahedron* **1997**, *53*, 5291–5304.
(49) Zhang, J.; Shi, Y. *Tetrahedron Lett.* **2000**, *41*, 8075–8078.
(50) Carpino, L. A.; Shroff, H.; Triolo, S. A.; Mansour, E.-S. M. E.; Wenschuh, H.; Albericio, F. *Tetrahedron Lett.* **1993**, *34*, 7829–7832.
(51) Barton, D. H. R.; Fontana, G.; Yang, Y. *Tetrahedron* **1996**, *52*, 2705–2716.

SCHEME 2



a) 20% piperidine/DMA; b) PyBOP, DIPEA, DMA; c) EDC or Mukaiyama reagent, DMF; d) 94:3:3 TFA:H_2O:TIS (triisopropylsilane); e) pentafluorophenyl chlorothioformate, DIPEA, CH_2Cl_2; f) $PbfNH_2$, potassium *t*-butoxide, DMSO.



solid phase is shown in Scheme 2. A protected amino acid molecule 4 was first attached to Rink Amide MBHA resin through standard peptide chemistry. After removal of the amine protection, the free amine was turned into the desired guanidine function through two routes. The first method is the direct guanidinylation with a Pbf-thiourea such as 2 (reaction sequence 1 in Scheme 2). This route produced the desired final N,N′-substituted guanidine in high overall yield (82% for 7a from resin-supported amine 5). However, this method is sometimes less attractive due to the requirement for pre-synthesis of Pbf-thioureas in solution. Alternatively, we demonstrated that one can form the Pbf-thiourea on solid support and subsequently produce the desired guanidine compounds. As shown by reaction sequence 2 in Scheme 2, the resin-supported amine 5 was first turned into pentafluorophenyl thiocarbamate. Reaction with PbfNHK, as in the solution synthesis, produced the desired Pbf-thiourea 6 on solid-support. Subsequent guanidinylation on solid support was smoothly carried out with a variety of amine nucleophiles in the presence of either EDC or the Mukaiyama reagent. Again, sterically hindered aliphatic amines or electron-deficient arylamines all gave very high overall isolated yields of the final guanidine compounds after TFA cleavage and HPLC purification.

Our solid-phase synthesis was also demonstrated with an alternative amino acid starting material in place of 4. Starting with an aromatic Fmoc-protected amino acid 8, compounds 9a–c were efficiently synthesized following either route in Scheme 2. We should note here that using α-amino acids as starting materials for the preparation of the corresponding guanidine derivatives will not lead to isolation of desired products under our reaction conditions since this class of compounds is known to go through intramolecular cyclizations.[24]

The fate of the side product due to the removal of Pbf protection by TFA treatment is also worth noting. TFA-based removal of Pbf from a guanidine group is expected to generate the corresponding aromatic ring moiety without trapping. Since we use TIS ($(i\text{-Pr})_3SiH$) as a scavenger, Pbf may also be trapped by TIS. In either case, however, none of these byproducts has significant solubility in aqueous solution while in contrast our desired guanidine products as their TFA salts do. A simple aqueous extraction and filtration removed most of these contaminants before the final HPLC purification (see the Supporting Information). Alternatively, washing the crude cleavage product with ethyl ether also served to remove most of the contaminants before HPLC purification.

With the success of either solution- or solid-phase guanidinylation using Pbf activated thiourea, it is worth comparing in more detail our method to other known procedures for preparing N,N′-substituted guanidines, with regard to our goals stated earlier. The foremost concern is the reaction efficiency of our approach. To this end, our methodology proved to be comparable to many other approaches. For example, in solution synthesis, the yields of guanidinylation (Scheme 1) are similar to solution phase guanidinylation using alkoxycarbonyl thioureas.[21] On solid-phase synthesis, our method produced slightly better isolated yields (Scheme 2) than syntheses performed with *N*-ethoxycarbonyl thiourea (63–83%).[9] Certainly, our method does not leave a guanidine protecting group after TFA cleavage, which is advantageous when water soluble guanidinium compounds are desired.

Compared to other synthetic methods using solid or soluble support for N,N′-substituted guanidines starting from two amines, our methodology also has its own merits. For example, Kearney et al.[29] and Chang et al.[25] each described a methodology through activation of thioureas by methylation to form isothioureas, while the Burgess group reported direct conversion of thioureas to carbodiimides through the use of Mukaiyama reagent.[19] These methods have the advantage that there is no need to handle protecting groups for the guanidine moiety during synthesis or cleavage. However, even though our method has to deal with the Pbf protecting group, its removal is achieved concomitantly with product cleavage and the side product was readily removed with simple aqueous extraction and filtration. On the other hand, the isothiourea method by Kearney et al. requires heating in sealed pressure-tube for 12 h at 80 °C for guanidinylation. In some instances, the final product purities were less than 50%, which may be attributed to a variety of

reasons including repeated treatment of the resin-bound materials with MeI for activation.[29] The isothiourea method described by Chang et al. also requires heating in DMSO, and the obtained yields vary widely (11−74%). The carbodiimide method described by the Burgess group is currently restricted to the starting amine being aniline derivatives bearing electron-withdrawing groups.[19] Compared to these three methods, our synthetic route can produce equivalent or better yields at room temperature for 12−18 h and is applicable to starting alkylamines.

There are two other syntheses of N,N′-substituted guanidines on solid-support, each using a carbonyl-based thiourea activation strategy.[24,28] The use of Wang resin to form TFA-labile alkoxycarbonyl activated thiourea allowed Josey at al to achieve efficient synthesis of guanidines with N,N′ and other types of substitution patterns.[28] However, the authors only specifically described product purity for phenyl and allyl based thioureas, and left out yield information for other starting alkyl thioureas. This makes it hard to compare with our current method. The method described by Wilson et al. used a novel carboxypolystyrene to produce acyl activated thiourea for guanidinylation. However, this method generally produced isolated yields in the 40% range,[24] while our method produced yields in the 70−95% range.

Compared to other guanidinylating methods based on TFA-sensitive arylsulfonyl groups (Mtr- and Pmc-based isothioureas,[36] and Mtr-based carbonimidodithioate,[39] Mtr: 4-methoxy-2,3,6-trimethylbenzenesulfonyl; Pmc: 2,2,5,7,8-pentamethylchroman-6-sulfonyl), our methodology can carry out guanidinylation at room temperature while the other methods invariantly require heating to displace the final MeS− group in the isothiourea moiety. Although we chose to use Pbf-based reagents, we expect that alternative use of either Mtr or Pmc should provide equal efficiency in reaction because the Pbf group contains a more electron-donating aromatic moiety than either Pmc or Mtr.[50] As for concomitant removal of these arylsulfonyl groups during final TFA cleavage, the Pbf or Pmc would probably perform equally in practical terms (half-life in TFA as guanidine protection: Pbf, 8 min;[50] Pmc, 13 min[50]), while the Mtr group should also be adequate since its complete removal can be achieved in one to 5 h,[52,53] which is compatible to general solid-phase cleavage protocols.

In summary, we have demonstrated the synthesis of N,N′-substituted guanidine compounds through TFA-sensitive arylsulfonyl activated thiourea that meets our aforementioned goals. The high reaction yield, mild reaction condition, and the flexibility of solution or solid-phase synthesis should make it a valuable extension to current synthetic methodologies. In addition, TFA-sensitive arylsulfonyl moieties are excellent guanidine protection groups. It is well-known that a single arylsulfonyl modification is sufficient to suppress side reactions on an arginine side chain in peptide chemistry, in contrast to carbonyl based protections.[54] This suggests that our novel guanidine synthesis method can be incorporated into other multiple-step synthetic procedures and provide a desirable protection for the guanidine moiety during the entire process. We are currently investigating such properties and will report our findings in due course.

Acknowledgment. We thank Dr. Wim Hol for stimulating discussions. This work was supported in part by the National Institutes of Health (GM94655, AI44954, and AI34501).

Supporting Information Available: Experimental procedures and characterization of compounds **2**, **3**, **7**, and **9**. This material is available free of charge via the Internet at http://pubs.acs.org.

JO026807M

(52) Fujino, M.; Wakimasu, M.; Kitada, E. *Chem. Pharm. Bull. Jpn.* **1981**, *29*, 2825.

(53) Atherton, E.; Sheppard, R. C.; Wade, J. D. *J. Chem. Soc., Chem. Commun.* **1983**, 1060−1062.

(54) Rink, H.; Sieber, P.; Raschdorf, F. *Tetrahedron Lett.* **1984**, *25*, 621−624.

VITA

Guangtao Zhang

University of Washington

2006

EDUCATION

Ph.D. in Organic Chemistry, 2006 Department of Chemistry, Biomolecular Structure Center (BMSD), University of Washington, Seattle, Washington.

Thesis: Design, Synthesis and Evaluation of Cholera Toxin Inhibitors and α-Helix Mimetics of Dormancy Survival Regulator.

Research Advisor: Erkang Fan

M. Sc. in Organic Chemistry, 2002 Department of Chemistry, University of Washington, Seattle, Washington.

B. Sc. in Chemistry, 2000 Department of Chemistry, University of Science and Technology of China, Hefei, Anhui, P. R. China.

RESEARCH INTERESTS

- Structure-based drug design
- X-ray crystallography
- Combinatorial chemistry and structure activity relationship
- Synthetic organic chemistry

Publications

1. Zhang G., Wisedchaisri G., Hol G. J. W., Fan E. Constraint Peptides as α-helix Mimetics to Disrupt the Protein-Protein Interaction in Dormacy Survival Regulator (Manuscript in preparation).

2. Zhang G., O'Neal C. J., Verlinde C. L. M. J., Hol G. J. W., Fan E. Design, Synthesis and Evaluation of Inhibitors for Cholera Toxin. *Abs. Pap. Am. Chem. Soc.* **2003**, conference poster.

3. Li J., Zhang G., Zhang Z., Fan E. TFA-Sensitive Arylsulfonylthiourea-Assisted Synthesis of N,N'-Substituted Guanidines. *J. Org. Chem.* **2003**, 68(4), 1611-1614.

4. Yin Y., Zhang G., Xia Y. Synthesis and Characterization of MgO Nanowires Through a Vapor-phase Precursor Method. *Adv. Func. Mater.* **2002**, 12(4), 293-298.

5. Ding Y., Zhao H., Sun Y., Zhang G., Wu H., Qian Y. Superstructured Magnesium Hydroxide Sulfate Hydrate Fibres. Photoluminescence study. *Int. J. Inorg. Mater.* **2001**, 3(2), 151-156.

6. Ding Y., Zhang G., Wu H., Hai B., Wang L., Qian, Y. Nanoscale Magnesium Hydroxide and Magnesium Oxide Powders: Control Over Size, Shape, and Structure via Hydrothermal Syntheses. *Chem. Mater.* **2001**, 13(2), 435-440.

7. Ding Y., Zhang G., Zhang S., Huang X., Yu W., Qian Y. Preparation and Characterization of Magnesium Hydroxide Sulfate Hydrate Whiskers. *Chem. Mater.* **2000**, 12(10), 2845-2852.