Expression and mutagenesis of recombinant cholera toxin A subunit
by Kirsten Louise Vadheim

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Microbiology
Montana State University

Abstract:
Vibrio cholerae causes a severe, potentially life-threatening diarrheal disease which is the result of the intoxication of epithelial cells of the small intestine by the cholera toxin. Cholera, pertussis and E. coli heat labile toxins are all protein exotoxins of the A/B format which exert their intoxicating effects within eucaryotic cells by ADP-ribosylating G proteins involved in the regulation of the adenylate cyclase complex, resulting in the increase of intracellular cAMP. These toxins share two short regions of sequence similarity within the enzymatically active A subunits. Substitution of the amino acid lysine for arginine at position nine in the first region of sequence similarity of pertussis toxin resulted in a protein with no detectable enzymatic activity. These results suggested that this first region of sequence similarity may be implicated in the ADP-ribosyltransferase activity of the toxin.

A previously poorly characterized gene coding for cholera toxin was subcloned, sequenced and expressed in E. coli. A series of site-specific mutations within the first region of sequence similarity was generated by polymerase chain reaction mutagenesis. Assays of the resulting proteins revealed that substitution of lysine for arginine, or arginine for aspartic acid, reduced ADP-ribosyltransferase activity to background levels, as did deletion of the entire region of eight amino acids. Substitution of a glycine for proline at position twelve had no effect on enzymatic activity, however.

Conservative substitutions of particular single amino acids within the first region of sequence similarity of cholera toxin resulted in reduction of enzymatic activity of the mutant cholera toxins to background levels. This corroborates previous studies with pertussis toxin and E. coli heat labile toxin. We conclude that the first region of sequence similarity in these toxins is important for enzymatic activity of the respective proteins, and await the crystallization of cholera toxin for a more detailed understanding of the relationship of these mutagenic studies to NAD binding sites within the toxin.

EXPRESSION AND MUTAGENESIS OF RECOMBINANT CHOLERA TOXIN A SUBUNIT

by

Kirsten Louise Vadheim

Co-advisors: Jerry M. Keith and Clifford W. Bond

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

of

Doctor of Philosophy

in

Microbiology

MONTANA STATE UNIVERSITY
Bozeman, Montana

June 1991

APPROVAL

of a thesis submitted by

Kirsten Louise Vadheim

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

June 12, 1991
Date

Co-Chairperson, Graduate Committee

June 14, 1991
Date

Co-Chairperson, Graduate Committee

Approved for the Major Department

June 14, 1991
Date

Head, Major Department

Approved for the College of Graduate Studies

June 27, 1991
Date

Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this thesis should be referred to University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute copies of the dissertation in and from microfilm and the right to reproduce and distribute by abstract in any format."

Signature Kirsten L Vadheim

Date June 14, 1991

This dissertation is dedicated to Marilyn and Roger Vadheim,
whose love and example made it possible.

TABLE OF CONTENTS

TABLE OF CONTENTS — Continued

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES -- Continued

ABSTRACT

Vibrio cholerae causes a severe, potentially life-threatening diarrheal disease which is the result of the intoxication of epithelial cells of the small intestine by the cholera toxin. Cholera, pertussis and *E. coli* heat labile toxins are all protein exotoxins of the A/B format which exert their intoxicating effects within eucaryotic cells by ADP-ribosylating G proteins involved in the regulation of the adenylate cyclase complex, resulting in the increase of intracellular cAMP. These toxins share two short regions of sequence similarity within the enzymatically active A subunits. Substitution of the amino acid lysine for arginine at position nine in the first region of sequence similarity of pertussis toxin resulted in a protein with no detectable enzymatic activity. These results suggested that this first region of sequence similarity may be implicated in the ADP-ribosyltransferase activity of the toxin.

A previously poorly characterized gene coding for cholera toxin was subcloned, sequenced and expressed in *E. coli*. A series of site-specific mutations within the first region of sequence similarity was generated by polymerase chain reaction mutagenesis. Assays of the resulting proteins revealed that substitution of lysine for arginine, or arginine for aspartic acid, reduced ADP-ribosyltransferase activity to background levels, as did deletion of the entire region of eight amino acids. Substitution of a glycine for proline at position twelve had no effect on enzymatic activity, however.

Conservative substitutions of particular single amino acids within the first region of sequence similarity of cholera toxin resulted in reduction of enzymatic activity of the mutant cholera toxins to background levels. This corroborates previous studies with pertussis toxin and *E. coli* heat labile toxin. We conclude that the first region of sequence similarity in these toxins is important for enzymatic activity of the respective proteins, and await the crystallization of cholera toxin for a more detailed understanding of the relationship of these mutagenic studies to NAD binding sites within the toxin.

INTRODUCTION

Cholera

History

"I am poured out like water, and all my bones are out of joint:
my heart is like wax; it is melted in the midst of my bowels.
My strength is dried up like a potsherd,
and my tongue cleaveth to my jaws;
And Thou hast brought me into the dust of death."

Psalm 22

Although there is no direct scientific evidence to support the idea, cholera may be one of the many diarrheal diseases that has plagued humanity throughout recorded history. It certainly has been endemic in the Indian subcontinent for centuries. Within the last two centuries seven pandemics have emerged from India, beginning in 1817 and ending with the most recent in 1961-1975. The second pandemic in about 1829 was the first to reach the New World, spreading throughout the whole of North and South America in addition to Russia, Europe and Great Britian. Repeated importations of the

disease brought about the third pandemic in the 1850's, the worst on record. Mortality rates were often 50 - 60% {van Heyningen and Seal, 1983}.

The first breakthroughs in understanding the causes and prevention of cholera occurred in the 1850's. Filippo Pacini reported the etiologic agent of cholera as a tiny curved bacterium, *Vibrio cholerae*, in 1854, and Robert Koch re-discovered it in 1883, naming it *Bacillus virgulus*. The British physician John Snow, by analyzing the water sources of London households, demonstrated that the disease was transmitted by sewage-contaminated water. His suggested preventive measures included personal cleanliness, avoidance of fecal contamination of food and drink, and destruction of the bacteria by cooking food and boiling water {van Heyningen and Seal, 1983, Holmgren and Svennerholm, 1983}.

The last major epidemic in the Western Hemisphere was in 1866-67. Cholera remains endemic in India, Pakistan and Bangladesh, with occasional flareups in these countries and throughout southern and southeast Asia and Africa, but the mortality rate has been reduced dramatically by the development of oral rehydration techniques. Diarrhea remains the leading cause of infant mortality in Third World countries but is more often due to rotavirus, *Escherichia coli* and dysentery bacilli than vibrios. However, in endemic areas cholera continues to cause significant morbidity, with a mortality rate usually less than 1%. Unfortunately, mortality rates still reach 60% in the initial phases of outbreaks in some areas, primarily due to inadequate medical care {van Heyningen and Seal, 1983, Isselbacher *et al.*, 1980, Siddique *et al.*, 1988}.

On January 29, 1991, reports of an outbreak of severe diarrhea in a coastal region of northern Peru reached medical authorities. *Vibrio cholerae* 01, Inaba, biotype El Tor, was isolated from patients' stools. Active surveillance, a national laboratory network and a public information campaign were immediately implemented. Within two weeks of the first reported case of diarrheal disease, 1,859 individuals had been hospitalized with severe gastroenteritis, and 66 died. {Centers for Disease Control, 1991} As of April 26, 1991, there had been 169,255 cases of probable cholera in Peru, 1,007 in Ecuador, 176 in Brazil, 26 in Chile, four in Brazil, and four imported cases in New Jersey {Pan American Health Organization, personal communication}. Clearly, cholera remains a very real threat to millions of people.

Clinical Disease

The most frightening aspects of cholera are its sudden onset and the rapidity with which its victims deteriorate. The incubation period is from six to forty-eight hours. Symptoms begin suddenly with vomiting and diarrhea, usually without any prodrome. Typical rice-water stools continue after the vomiting has ceased, resulting in the loss of 20 to 30 liters of water per day in severe cases. This extreme and rapid dehydration produces the typical picture of cholera: cold, clammy skin, a feeble and often imperceptible pulse, tachypnea, pinched face, sunken eyes, poor skin turgor, and shrivelling of the skin on the hands and feet. The patient is cyanotic, hypotensive and xerostomic. If untreated, severely afflicted patients may die within a few hours of the onset of symptoms. This sudden transformation of a seemingly

normal, healthy individual to an enfeebled, cadaverish victim within a few hours is undoubtedly responsible for the panic that often accompanied the cholera pandemics. The disease usually runs its course in two to seven days, and significant sequelae are rare {van Heyningen and Seal, 1983, Collier and Mekalanos, 1980, and Isselbacher *et al.*, 1980}.

Oral rehydration of a cholera patient was first attempted by a Scottish physician, Thomas Leith, in 1832. He was able to revive the individual, but once rehydrated the patient simply continued to purge and quickly died {van Heyningen and Seal, 1983}. Not until the 1960's was a simple, effective formula for oral rehydration developed. The standard recipe for one liter of rehydration fluid is 20 g glucose, 4.2 g NaCl, 4 g $NaHCO_3$, and 1.8 g KCl.

Treatment usually consists of re-establishing the electrolyte balance by administration of intravenous salt solutions and encouraging the patient to drink the rehydration fluid to replace water lost through diarrhea. Streptomycin or tetracycline will decrease the volume of stool released and may shorten the course of disease by interrupting the multiplication of the bacteria, but antibiotics are of little benefit without simultaneous fluid replacement {Holmgren and Svennerholm, 1983}. Mass chemoprophylaxis, vaccination and quarantine of cholera victims have been proven to be ineffective, and simply divert scarce resources from efforts to adequately treat victims and control further spread of cholera {Centers for Disease Control, 1991}.

Bacteriology and Epidemiology of *Vibrio cholerae*

Vibrio cholerae is an aerobic, Gram negative, curved rod with a single polar flagellum which is responsible for the bacterium's rapid motility {Brock and Madigan, 1988}. In addition to the well-defined cholera toxin (Ctx) proteins, *V. cholerae* elaborates a considerable number of virulence factors, including hemolysins, a neuraminidase, a variety of proteases, hemagglutinins, at least two types of fimbriae, a Shiga-like toxin, and several outer membrane proteins that are regulated along with other known virulence factors and which may be involved in osmoregulation {Holmgren and Svennerholm, 1983, and Taylor, 1989}.

Both O and H somatic antigens are present on *V. cholerae*. All cholera epidemics traced to date have been caused by serogroup O type 1 (O1) bacterial strains. Three antigenic factors, A, B, and C, are used to subdivide the O1 serogroup into the serotypes, shown in Table 1.

Serotype	O Factors
Ogawa	AB
Inaba	AC

Table 1. Serogroup division of cholera vibrios.

Serologic conversion among the serotypes can occur during natural infection; this seems to be related to the appearance of agglutinating antibodies in the serum {Zwadyk *et al.*, 1980}.

Serogroup O1 contains the biotypes cholerae (classical) and El Tor {Zwadyk *et al.*, 1980}. Classical *V. cholerae* was responsible for the first six

pandemics and El Tor vibrios for the seventh. Although both can cause disease, classical strains generally synthesize more toxin {Miller and Mekalanos, 1985}. The two biotypes are differentiated on the basis of physiological properties such as hemolysin production and polymyxin sensitivity {van Heyningen and Seal, 1980}.

The El Tor biotype was named after a quarantine station on the Sinai Peninsula where this strain was first discovered in 1889. El Tor frequently infects without causing disease, and was discovered in pilgrims who had died without any symptoms of cholera. It has a much greater capacity to survive in the environment than classical strains, and has been isolated from aquatic environments and shown to grow in many foods. Because the ratio of inapparent infections to actual cases may be as high as 100 to 1, El Tor can quickly spread within a population and persist easier than the classical vibrios. When the El Tor biotype first appeared in the Celebes, it caused a cholera-like disease with a low case-infection rate but a fatality rate of 50 to 60%, at least as high as classical cholera {van Heyningen and Seal, 1983}.

Nonagglutinable (NAG) or non-cholera vibrios (NCVs) are *V. cholerae* serotypes other than O1. At one time these non-O1 strains were thought to be non-pathogenic, but NAG vibrios are now known to be responsible for up to 5% of the acute diarrheal illness in cholera-endemic areas. Nearly 50 different non-O1 serotypes have been reported so far {Kaper *et al.*, 1981, and Isselbacher *et al.*, 1981}.

All age groups are susceptible to cholera when it spreads into new areas, but in endemic regions it is primarily a disease of childhood. This is demonstrated by the close inverse relationship between the attack rate of

cholera and age. In addition, serum vibriocidal antibody titers increase with the age of individuals in endemic areas {Holmgren and Svennerholm, 1983}.

The primary route of transmission of cholera is through fecally-contaminated drinking water. Because there are a number of significant secondary routes of infection, improvement of drinking water quality alone does not effect major reductions in cholera incidence in endemic areas. Other routes of infection include ingestion of water during bathing, eating contaminated food, and interpersonal contact {Briscoe, 1984}.

Natural reservoirs for cholera outside of endemic regions are poorly understood. A chronic gall bladder carrier state has been observed in elderly convalescent cholera patients, but its relevance to the spread of the disease remains speculative {Isselbacher *et al.* 1980 and 1981}. Comparison of strains involved in a limited epidemic in Louisiana in 1978 with a case in Texas in 1973 and an imported case from Mexico in 1983 showed all three to be identical: biotype El Tor, serotype Inaba, strongly hemolytic, with unique phage sensitivity patterns and with identical *ctx* gene sequences different from those of the pandemic strain being isolated from the rest of the world at that time. This suggests that there may be a strain of *V. cholerae* which has persisted in the U.S. coastal waters for many years, perhaps establishing its own free-living cycle in water {Blake *et al.*, 1983}.

Cholera Toxin

Structure

As early as the 1880s Robert Koch suggested that cholera vibrios produce a toxin which was responsible for the disease. Not until 1959, however, was the existence of an enterotoxin proven and its importance in disease demonstrated {De, 1959, and Holmgren and Svennerholm, 1983}.

Ctx, also known as choleragen {Finkelstein *et al.*, 1964} is a secreted, heat labile protein with a molecular weight of 84,490 {Pearson and Mekalanos, 1982, and Lockman *et al.*, 1984}. This toxin is one of several A/B model toxins where A represents the enzymatically active subunit and B represents the binding subunit. Its subunit composition is AB_5. The *ctx* genes are arranged in a single transcriptional unit with one promoter controlling expression of both subunits (see Figure 1). Some strains, particularly the classical strains, contain multiple copies of the operon in the *V. cholerae* chromosome (Mekalanos 1983).

Cholera Toxin Operon

Figure 1. Genetic organization of the cholera toxin operon.

The A subunit is synthesized as a single polypeptide with a molecular mass of 29,000 kilodaltons (kDa), but is enzymatically active in a proteolytically nicked form which has two disulfide-linked chains, A_1 (23 kDa) and A_2 (6 kDa) {Mekalanos *et al.*, 1983 and Collier and Mekalanos, 1980}. Cleavage occurs at a trypsin-sensitive site immediately preceding two adjacent serine residues, a dipeptide that is unique in the translational product of the *ctx* operon. Although the disulfide bond linking the two chains is easily cleaved by reducing agents, non-covalent forces tend to prevent dissociation of the two subunits under non-denaturing conditions {Lockman *et al.*, 1984}.

The primary translation product of *ctxA* is a 258 amino acid polypeptide with an 18 amino acid hydrophobic signal sequence. The B peptide is composed of 124 amino acids, also has an 18 residue leader sequence and contains one intrachain disulfide bond. The B oligomer (B_5) is highly resistant to dissociation by reducing agents {Mekalanos *et al.*, 1983, Collier and Mekalanos, 1980 and Lockman *et al.*, 1984}.

The *ctx* operon has a characteristic ρ-dependent termination signal. A twenty-five base region of dyad symmetry at the end of the B cistron contains a G + C-rich stem loop region and a poly(T) tail {Mekalanos *et al.*, 1983}.

Regulation

Both *ctxA* and *ctxB* are transcribed from a single promoter preceding the *ctxA* gene, resulting in a polycistronic mRNA transcriptional product which contains one copy each of the *ctxA* and *ctxB* genes. Disparate translational efficiencies appear to account for the increased production of CtxB to achieve a final ratio of five CtxB subunits to one CtxA. Mutations of

the *ctx* operon which placed the *ctxB* product under the control of the *ctxA* gene translational signals resulted in approximately nine-fold less CtxB than usual {Mekalanos *et al.*, 1983}. Comparison of the Shine-Delgarno sequences of *ctxA* and *ctxB* with an *E. coli* consensus sequence also indicated that the *ctxB* sequence is almost identical with the consensus sequence, and therefore possibly translates at a higher efficiency than from the *ctxA* sequence, which differs in several locations {Lockman *et al.*, 1984}.

The *ctx* genes lie in a genetic element with a structure very similar to that of transposons (Figure 2). A 2.7 kilobase (kb) repetitive sequence, RS1, is located adjacent to and upstream of the 4.3 kb core region which contains *ctxAB* and other genes involved in the regulation of cholera toxin {Mekalanos, 1983, Miller and Mekalanos, 1984, DiRita and Mekalanos, 1989}. RS1 also forms the junction between tandem duplications of the *ctx* genetic element and is sometimes found downstream as well, so that the RS1 direct repeats flank the core sequence. RS1 can transpose and is involved in a *recA*-dependent recombination process leading to the duplication or further amplification of tandem repeats of the *ctx* element {Goldberg and Mekalanos, 1986}.

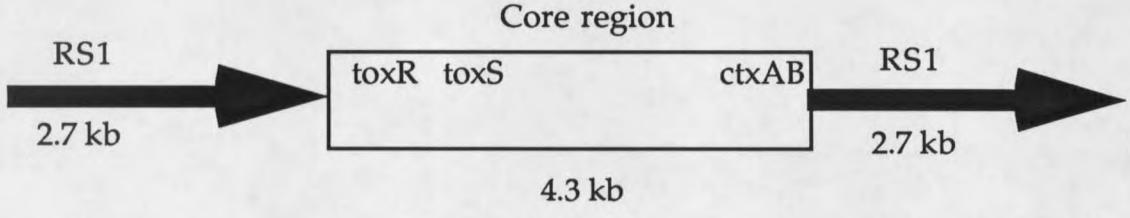

Figure 2. Schematic diagram of the ToxR regulatory region.

Ctx is not produced constitutively in *Vibrio cholerae* but is transcribed when activated by ToxR, a global regulatory protein that is the primary component in the transduction of environmental signals that lead to virulence gene expression {DiRita and Mekalanos, 1989}. The *toxR* gene, which is located upstream from *ctxAB* but within the same core element, encodes a protein of 32.5 kDa {Miller and Mekalanos, 1988}. ToxR is a transmembrane protein with cytoplasmic N-terminal and periplasmic C-terminal regions. The C-terminal domain is responsive to environmental signals such as changes in osmolarity. The N-terminus binds to the DNA sequence TTTTGAT found in multiple, tandem repeats in the region directly upstream from the *ctx* operon, thereby activating transcription from the *ctxAB* promoter. Other genes regulated by ToxR include the cluster *tcpABCDEFG*, which codes for a pilus that is the primary colonization factor of *V. cholerae*, the *acf* operon (*acfABCD*) that codes for an accessory colonization factor, and the *ompT* and *ompU* genes {V. Miller *et al.*, 1989}.

A second regulatory protein, ToxS, is the product of the *toxS* gene, located immediately downstream of the *toxR* gene in the core region of the *ctx* genetic element. The *toxR* gene was originally cloned from the classical Inaba strain 569B, which is a hypertoxinogenic but generally less virulent strain of *V. cholerae*. The 569B chromosome carrries a 1.2 kb deletion in the region where *toxS* is located. ToxS is a 19 kDa transmembrane protein with most of its sequence in the periplasmic space {V. Miller *et al.*, 1989}. It interacts with the C-terminal periplasmic domain of ToxR to confer the $ToxR^+$ state, probably by stabilizing a dimerized form of the ToxR protein {DiRita and Mekalanos, 1991}. ToxR and ToxS do not appear to have the

sensor-regulator relationship found in many other two-component bacterial virulence regulators {J. Miller *et al.*, 1989}.

A third regulatory gene involved in controlling expression of the *V. cholerae* virulence genes, *toxT*, may be a transcriptional activator whose promoter is activated by ToxR. ToxT can activate gene fusions that are dependent on *toxR* in *V. cholerae* but that cannot be activated in *E. coli* by cloned *toxR* alone {DiRita and Mekalanos, 1989}.

Pathophysiological Effects of Ctx

The human small intestine is essentially a hollow tube whose inner surface is covered with a continuous layer of columnar epithelium. This epithelium forms the outer layer of the gut mucosa and is organized into crypts (glands of Lieberkuhn), which penetrate the underlying lamina propria, and villi, which project into the lumen of the gut. Cells of the lower half of the crypts constantly proliferate and gradually migrate as they mature to the tips of the villi, where they are desquamated individually. Because of the rapid rate of cell proliferation and shedding, the entire luminal surface of the intestine is replaced approximately every five days {Junqueria and Carneiro, 1983}.

V. cholerae colonizes, but does not invade, the epithelium of the small intestine and secretes Ctx. The B protomer of the holotoxin binds to the galactosyl - N- acetylgalactosaminyl -(N-acetylneuraminyl) -galactosylglucosyl-ceramide (GM_1) ganglioside on the surface of the gut mucosal cells. GM_1 is comprised of a lipid moiety, ceramide, which is inserted into the lipid matrix of the cell membrane, and an oligosaccharide moiety which is exposed on the