Bacterial Conjugation and its Inhibition:

The Hows and Whys of Conjugation and What Can be Done to Control It

Scott A. Lujan

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the department of Biochemistry and Biophysics.

Chapel Hill, NC
2008

Approved by

Professor Matthew Redinbo, Ph.D.
Professor Richard Wolfenden, Ph.D.
Professor Steve Matson, Ph.D.
Professor Gary Pielak, Ph.D.
Professor Brenda Temple, Ph.D.

Abstract

SCOTT A. LUJAN: Bacterial Conjugation and its Inhibition: The Hows and Whys of Conjugation and What Can be Done to Control It

(Under the direction of Matthew R. Redinbo, Ph.D.)

Conjugation is the primary vehicle for the horizontal transfer virulence factor genes, such as antibiotic resistance, within and between bacterial strains. In certain epicenters, such as hospitals in less developed parts of the world, immune-compromised patients and misuse of antibiotics combine to select for the development and dissemination of these pathogenicity factors via conjugation. Inhibition of conjugation would prove a boon for curbing the creation and spread of new virulent or multi-drug resistant strains. DNA relaxases are the keystone proteins of each conjugative system. TraI is the relaxase of the F plasmid, the archetypal model system for conjugation. Toward revelation and inhibition of relaxase mechanisms, I used bioinformatics and limited proteolysis to find and identify new domains on the F TraI enzyme. I then solved TraI/DNA co-crystal structures that showed a novel DNA binding mode. Based on structure comparisons, sequence conservation, and mutant activity studies, I proposed a mechanism for TraI activity that required the existence of a dual phosphotyrosine intermediate, which has since been observed. I then demonstrated that bisphosphonate compounds mimicking this intermediate are nanomolar *in vitro*

TraI inhibitors. I determined the relaxase structure in complex with one such inhibitor, bound as predicted. We showed that several of these compounds are potent *in cell* inhibitors of conjugation that often selectively kill conjugation-capable cells, a novel antibiotic paradigm. We showed that oral treatment of gnotobiotic mice with two inhibitors, clodronate and etidronate, decreased the gastrointestinal F^+ bacterial load twentyfold without apparent side effects. Beyond the model system, etidronate showed selective *in cell* potency versus cells harboring a known clinical resistance-bearing R100 plasmid. Etidronate and clodronate are already clinically approved for the treatment of bone loss in humans. Toward more general relaxase inhibition phylogenetic analyses and studies with other medically relevant relaxases, including viral replicative relaxases, are ongoing.

Dedication

For my parents, who got me to this point,

and for Nan, who helped me see it through.

Acknowledgements

I would like to thank Rebekah Potts, Heather Ragonese, and Dr. Laura Guogas for all of their help in the execution of my crazy schemes. I would also like to thank Sarah Kennedy, Debi Haisch, Dhruthi Patel, Heather Bethea, and Lori Hannula for their assistance on the bench in recent years and Drs. Sompop Bencharit and Diem-Thu Lesher for their assistance and training early on. Joseph Lomino, Dr. Eric Ortlund, and Dr. Mike Miley, you listened, aided, and occasionally abetted, for which you have my appreciation. Thanks to Maureen Bower and the staff of the UNC Gnotobiotic Mouse facility, under the auspices of the Center for Gastrointestinal Biology and Disease, for the care and handling of our rodent subjects. Likewise, thanks to Dr. Balfour Sartor for mice, expertise, and facility time. My distinguished committee, Drs. Richard Wolfenden, Jan Hermans, Brenda Temple, Gary Pielak, and Steve Matson, thank you for your support and advice. Steve and Gary, you listened and advised as if I was a member of your own labs, for which I am grateful. Finally, to Dr. Matthew Redinbo, every time I asked for a little more rope, you gave it to me. You risked time, funding, and stress on me and on this project. I hope someday to be to be as good a mentor as you have been to me. I shall always be in your debt.

Table of Contents

List of Tables

List of Figures

1 Background and Significance

In many underdeveloped areas around the globe, antibiotic resistance among virulent bacterial strains is reaching crisis proportions. Bacterial conjugation is a major player in the spread of antibiotic resistance within and between bacterial strains. The F-plasmid conjugation system is used as a model system for general bacterial conjugation. TraI, also known as DNA Helicase I, is a keystone protein for F-plasmid conjugation: it possesses both helicase and relaxase activities necessary for the plasmid mobilization phase of conjugation (1). Transfer assays show that certain TraI truncations maintain nicking and unwinding activity but abrogate strand transfer (1). This suggests that TraI plays a crucial role beyond the mobilization phase. The mechanisms of TraI activity are not well understood, nor are its extended roles. F TraI relaxase mechanisms have been postulated based on the functions of TraI homologues, including TrwC of the R388 IncW plasmid and TraI of the RP4 IncP plasmid. However, architectural and biochemical data suggest that F TraI relaxase mechanisms vary somewhat from those of TrwC and a great deal from those of RP4 TraI (2, 3). Computational studies now suggest structural reasons for this difference and illustrate the relationship between structural details and related mechanisms. My

work will test these suggestions and reveal the structural basis for TraI function as a model for R-plasmid relaxase/helicases.

1.1 Humanity at War

Horizontal transfer of genetic information among bacteria is accomplished mainly through conjugation (4, 5). It is usually mediated by a conjugative plasmid residing in the cytoplasm or integrated into the genomic DNA of the donor cell. There is an energy cost associated with hosting a plasmid: plasmid repair, translation of plasmid proteins, and copy number maintenance all use valuable cellular resources. As with laboratory strains transformed with artificial vectors, daughter cells that lack foreign plasmids are at a selective advantage versus those that are hosts. This helps curb the spread of plasmid borne antibiotic resistance and virulence factors. Problems arise when a bacterial population, some individuals of which are plasmid hosts, is subjected to selective pressure, such as antibiotic treatments. The selective balance then shifts to daughter cells that inherit plasmid copies. This in itself is not a major concern for healthcare providers, as the modern suite of antibiotics allows some flexibility in circumventing resistance. However, continued selective pressure eventually encourages the development of new forms of resistance, which can then be disseminated via conjugative plasmids (4, 6). It is then probable that some bacteria will gain multiple forms of resistance, sometimes to entire classes of drugs (4). These multi-drug resistant (MDR) strains can become flexible enough to thwart antibiotic treatment altogether, becoming so-called “super bugs”. In some underdeveloped countries, the situation can result in local health crises, as

is the case with MDR strains of *mycobacterium tuberculosis* in Russia, Peru, certain African nations, and parts of China (7, 8). Resistance-bearing conjugative plasmids have been found in many of mankind's historical nemeses, including *Haemophilus influenzae* (bacterial meningitis), *Salmonella typhi* (typhoid fever), *Vibrio* strains (cholera and its relatives), and *Yersenia pestis* (plague) (6). Similarly, previously avirulent strains may be pathogenized by the introduction of virulence factors via conjugation. Such a process has been implicated by the localization of *Salmonella typhimurium* derived virulence genes on an F-like plasmid in pathogenic *E. coli* O157:H7 and by F-plasmid transfer genes on the pFra virulence plasmid of *Yersenia pestis* strain G8786 (9, 10). Medically, the potential benefits of understanding and fighting antibiotic resistance and virulence factors spread via conjugative plasmids cannot be overestimated.

1.2 Conjugation

Parallels have been drawn between conjugative mobilization and the rolling circle form of replication seen in many viruses. Indeed, all of the relaxases discussed herein are considered members of the Mob class of nickases, which falls (along with the Rep class) into the Rolling Circle Replication (RCR) superfamily of magnesium dependent nucleases (11). Though any true evolutionary linkage is unclear, conjugative plasmids, like viruses, are subject to selective pressure. Thus, many conjugative systems include genes that provide selective advantages to host cells. In *E. coli*, the R series of plasmids is particularly noteworthy for containing and transferring genes that confer antibiotic resistance on their hosts (4, 12, 13). Conjugative plasmids are often classified

into incompatibility groups, such as IncF for F-plasmid and R100, IncP for RP4 and RK2, IncW for R388, and Ti-type IncN for R46 (6). Most R plasmids also code for a repressor that limits their conjugation rate. The F-plasmid is similar to the rest of the R series - nearly identical to plasmid R100 - with the notable exception that the repressor gene of F lacks a start codon (14). F may thus be considered a constitutively active R100 mutant. As such, I will often refer to F and R systems interchangeably. Wild type F also lacks resistance or virulence, a rare example of a system that persists on efficiency alone, a true parasite conferring no selective advantage to its hosts. F-plasmid must thus rely solely upon its conjugation proficiency to propagate successfully. Among the most well characterized of all bacterial conjugation systems, The F-plasmid system is an ideal model system. ***My long term goal was to characterize the mechanism of plasmid to identify methods to inhibit the propagation of antibiotic resistance and bacterial virulence factors.***

There are four major phases in bacterial conjugation (Figure 1):

1. **Contact**: a physical connection is made between the donor and recipient cells, often through sex pili. The connection must span the cellular membranes and cell walls of both cells in order to create a conduit between opposing cytoplasmic spaces (12).
2. **Mobilization**: the conjugative plasmid is prepared for transfer. The plasmid is nicked and separated into component strands, which requires unwinding of the duplex DNA. Mobilization results in a single-stranded circular non-transfer strand and either a linear or circular single-stranded transfer strand (12).

3. **Transfer**: one linear strand is moved to the cellular membrane and passed into the intercellular conduit. Some systems include a “pilot” protein that guides the strand through the conduit to the recipient cytoplasm (15, 16).
4. **Replication**: the transferred strand is replicated and either re-circularized or integrated into the recipient genome. The non-transferred strand is also replicated, possibly in conjunction with Step 2 unwinding (12).

Specific protein systems are involved in each phase of conjugation, usually in the form of large, dynamic complexes. Pilus proteins initiate contact with the recipient cell and may help conduct the transferred strand across intercellular space. Adapter proteins are important for strand transfer, forming the link between extracellular pilus assemblies and the cytoplasm of the donor cell (12). Mobilization is conducted by a complex dubbed the relaxosome. The relaxosome is responsible for the initiation, execution, and termination of strand transfer. In the F-plasmid, the relaxosome is composed of three essential proteins (TraI, TraY, IHF) and a fourth (TraM) required for maximal activity (17-19).

1.3 TraI: Keystone of F Plasmid Conjugation

TraI, also known as DNA Helicase I, possesses both relaxase (transesterase) and helicase activities. TraI performs the initial nick that begins plasmid transfer, unwinds and separates the nucleotide strands, and has other proposed activities during both the mobilization and transfer phases of conjugation (20). This makes TraI a key protein in F-plasmid conjugation and thus essential to an understanding of conjugation as a whole.

1.4 All in the Family: The Relaxase Domain

The details of the diverse Mob family relaxase mechanisms are poorly understood. Multiple sequence alignments show that some relaxase/transesterase/nickases possess one catalytic tyrosine residue (TraA and VirD2 of *Rhizobia* and *Agrobacteria* pTi plasmids; TraI of plasmid RP4 IncQ) while other have a constellation of conserved tyrosines (TraA of *Corynebacteria* plasmids; TrwC of plasmid R388 IncW; TraI of plasmid R100 IncF plasmids). The IncF/W constellation of two tyrosine pairs follows a pattern of Yy-X_{3-26}-YY, where 'y' denotes either a tyrosine or a phenylalanine in some IncW relaxases (21). (Fig. 2) The tyrosine pairs are thought to be responsible for alternating transesterifications during sequential initiations/terminations of the mobilization phase in a rolling-circle manner as described for bacteriophage phi X174 (11, 12, 22). During the nicking (transesterification) reaction, a catalytic tyrosine forms a covalent bond with a phosphate of the nicked strand (1, 13). It has been postulated that R388 TrwC exploits this covalent bond and acts a pilot protein that guides the transferred strand from donor to recipient (12, 15). This was suggested through analogy with known pilot proteins of *Agrobacterium* (Ti-system) and RP4 TraI (RP4). However, experimental evidence is against an F-plasmid pilot, and at 1756 residues in length and 191 kDa in mass, TraI is far larger than known pilot proteins, an unlikely candidate for transmembrane secretion (15, 16). Mutant data with R388 TrwC show that all four tyrosines show varying degrees of capability for creating the initial nick (3). The exact reason for conserving redundant catalytic residues is unclear.

Mechanisms for TraI relaxase activity have been posited in analogy with R388 TrwC, for which an abundance of mutant activity data exist (3). More recent mutational data on TraI activity appear to contradict many of the TrwC results, suggesting that either the mechanisms differ or that a much more complicated story is in the offing (2). Crystallographic data exist for apo TraI (23) and DNA bound TrwC (24), but both structures leave much to be desired. The TraI structure lacks the thirty residues of the DNA binding loop (disordered), and the TrwC structures lack residues and DNA proximal to the active site (disordered and not present, respectively).

1.5 Something New has been Added: The TFG Domain

The central 600 residues of F TraI (situated between the relaxase and helicase domains) and the carboxyl terminal 300 residues have long been a mystery. Transfer is lost upon their truncation or deletion, and activity is not recovered upon reconstitution of the helicase and relaxase domains (1). These regions lack any previously defined domain architecture and there are no homologous regions in IncP, Ti, or even the closely related IncW incompatibility groups (R388 TrwC consists of a relaxase domain and a helicase domain connected by a short linker with no C-terminal extension) (1). I have now shown, via computational analyses, that there resides in TraI residues 300 to 600 a Rossman fold domain, similar in structure to the GTPase domain of *Thermus aquaticus* Ffh, dubbed the TraI Ffh GTPase-like domain (TFG) (25). This seeming incongruity actually fills in an essential gap in the conjugation sequence of events. Consider this question: if mobilization occurs in the cytoplasm, how

does the transferred strand move to the inner membrane and couple with the conjugative secretion system? The answer hinges on an understanding of the function of Ffh.

Ffh is a member of the bacterial Signal Recognition Particle (SRP). The SRP is charged with the task of translocating ribosomes to the inner membrane upon translation of membrane targeting sequences (26). Ffh has two domains: the M domain, in SRP parlance, which recognizes and binds targeting sequences and a small SRP RNA; and an N/G domain, consisting of an N-terminal four-helix bundle (N) and a C-terminal GTPase subdomain (G). After signal recognition and binding, the complex (Ffh, the ribosome, tRNA, nascent polypeptide, Ffh bound GTP, and a small SRP RNA) translocates to the inner membrane, where the Ffh N/G domain complexes with an adapter protein, FtsY. FtsY also consists of two domains: a C-terminal membrane associating domain; and an N-terminal N/G domain. The N/G domains interact along a large interface that includes both GTPase active sites. After the nascent polypeptide (signal sequence) passes to the secretion system (membrane pore), simultaneous GTP hydrolysis occurs, mediating Ffh/FtsY dissociation (27, 28).

The SRP system suggests a novel activity for TraI only if an FtsY-type adapter protein exists in the IncF. TraD is an F encoded hexameric membrane protein that interacts with TraI, non-specific DNA, and the conjugal secretion system. TraD has a short transmembrane portion and a large Rossman type domain (with a subset of canonical NTPase sequences), capped by a four-helix bundle. Though it has no native NTPase activity, the TraD p-loop (putative active

site) faces out into the cytoplasm (based on homology modeling versus R388 TrwB (29, 30)). This makes TraD an excellent candidate to play FtsY to TraI's Ffh. (Fig. 3).

1.6 Summary

My scientific priorities have shifted since my original thesis proposal. A crystal structure of the TFG domain would show unequivocally whether or not this domain is structurally homologous to the FtsY family of small GTPases (Aim 2). Likewise, a co-crystal structure of TraI/TraD or TraI/TraD/DNA, would explore a novel TraI function in the transfer phase of conjugation (Aim 3). However, these aims are left to future work, as work on Aim 1 (below) and the original long-term goal of conjugation inhibition proved to be wildly more successful than originally anticipated.

An understanding of bacterial conjugation is essential for humanity's long war with our microbial adversaries. The F-plasmid conjugation system is an ideal model for a large family of medically relevant conjugative plasmids. As the keystone of the F-plasmid system, structural knowledge of the various domains of TraI has been invaluable in the development of weapons in our continued fight.

I originally proposed (Aim 1) to acquire structural data showing the relative partial occupancies of the four F TraI tyrosines coordinated to either small oligonucleotides or intermediate/transition state analogues in order to shed light on the mechanism. I solved the structure of the TraI relaxase domain in complex with one DNA nucleotide. Comparison with other another, apoTraI structure (23)

and structures of the R388 homologue, TrwC (24), allowed me to propose a multi-tyrosine relaxase mechanism (31). This in turn allowed me to find small molecule inhibitors of relaxase activity, which I showed to be potent both *in vitro* and *in cell.* I then solved the structure of the relaxase domain in complex with DNA and one of these inhibitors, bound as predicted. I then showed that some of these inhibitors, including two that are already approved for human use to treat unrelated conditions, selectively kill donor cells both *in vitro* and in a gnotobiotic mouse gut. Studies with plasmid systems closely related to the F plasmid have been promising. Should inhibition prove viable for broad host range plasmids and viruses with replicative relaxases, and should human safety and drug potency prove acceptable, the implications for human health are profound.

1.7 Figure Legends

Figure 1. Bacterial Conjugation

In a simplified schematic of bacterial conjugation, the left-hand cell is R+(the donor), the right-hand cell is initially R^- (the recipient).

A. **Pre- conjugation**: the donor has plasmid encoded drug resistance, the recipient has genomic drug resistance.
B. **Contact:** a conduit forms for DNA transfer; a mating signal prompts relaxosome formation on the plasmid.
C. **Mobilization**: the plasmid is nicked, unwound, transported to the membrane, and replicated.
D. **Transfer**: the nicked strand is transported across the conduit and into the recipient.
E. **Replication**: the conduit is terminated and the transferred strand is either circularized (religated) and replicated (as shown) or integrated into the recipient genome (not shown).
F. **Expression**: Plasmid genes, including those for new antibiotic resistance, are expressed in the recipient, now R^+.

Figure 2. Relaxase/Nickase Helicases.

A representative set of proteins containing relaxase and helicase domains. Relaxases in which these are not coupled (such as IncP RP4 TraI) are not shown. Groups **a** and **b** represent the relaxase proteins common to R-class

bacterial conjugative plasmids from related incompatibility (Inc) groups. Color variations within a domain indicate approximate sub-domains.

A. These F-type proteins have a relaxase domain with four catalytic tyrosines (Yx4), a proposed Ffh GTPase-like domain (TFG), and a helicase domain with a characteristic DESS Walker B motif.

B. IncW relaxases have a relaxase domain with three or four catalytic tyrosines and a helicase domain with a characteristic DEAX Walker B motif.

C. Ti-type relaxases (involved in bacteria-to-plant conjugation) have a smaller helicase domain with an R388-like DEAX Walker B motif. *Agrobacterium* and *Rhizobium* relaxases have one catalytic tyrosine (IncP-like); *Corynebacteria* bear an IncF-like four-tyrosine motif.

Figure 3. Models of Proposed TraI/TraD Interaction

A. TraI and TraD during plasmid mobilization. Hexameric TraD is represented in **grey**. TraI domains colored as in Fig. 2: relaxase **red**, TGF **blue**, and helicase **green**. TraI helicase action pumps the nicked DNA strand (**cyan**) through the Type IV Secretion System (T4SS) pore complex in the direction of the blue arrow.

B. A homology model of the TGF/TraD interaction. The TGF domain (**cyan**) is modeled against Ffh. TraD (**blue**) is modeled against TrwB. Putative NTP binding loops are **red**. On the right, a second monomer in the TraD hexamer is shown as a **grey** silhouette.

Figure 1-1 Bacterial Conjugation

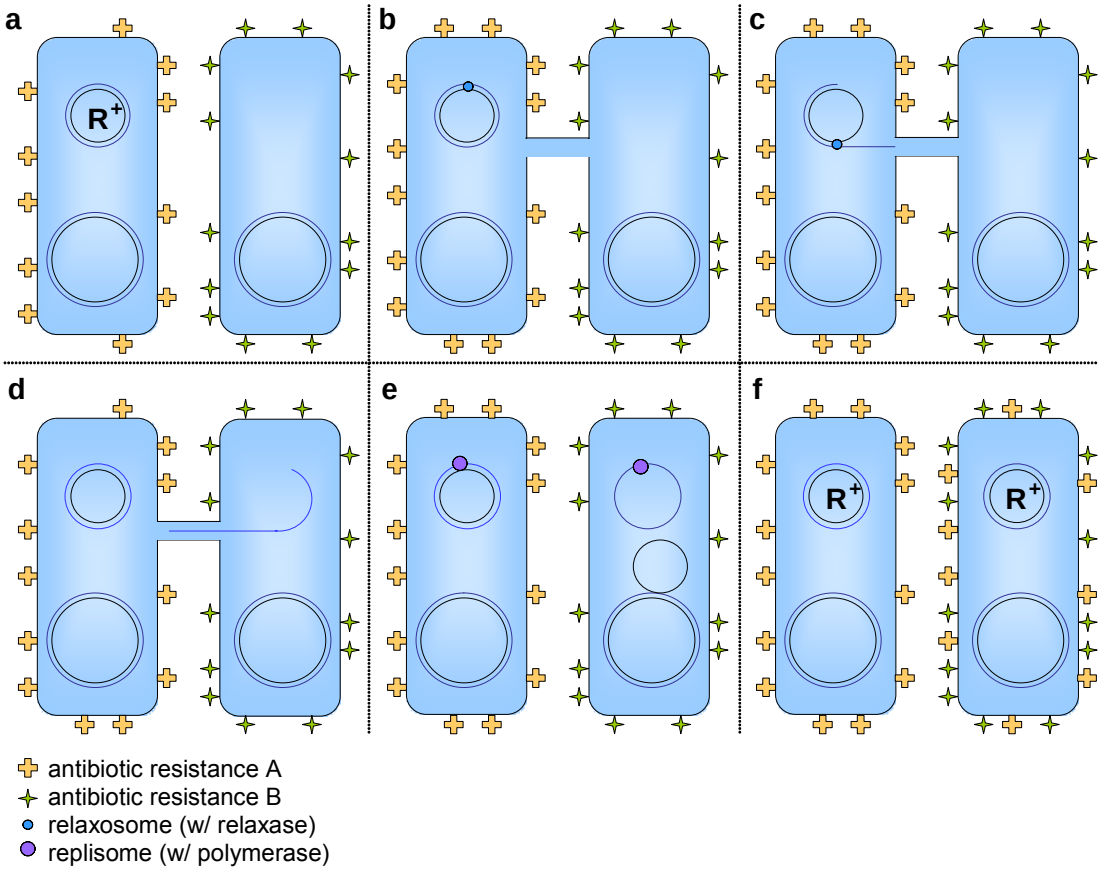


Figure 1-2 Relaxase/Nickase Helicases

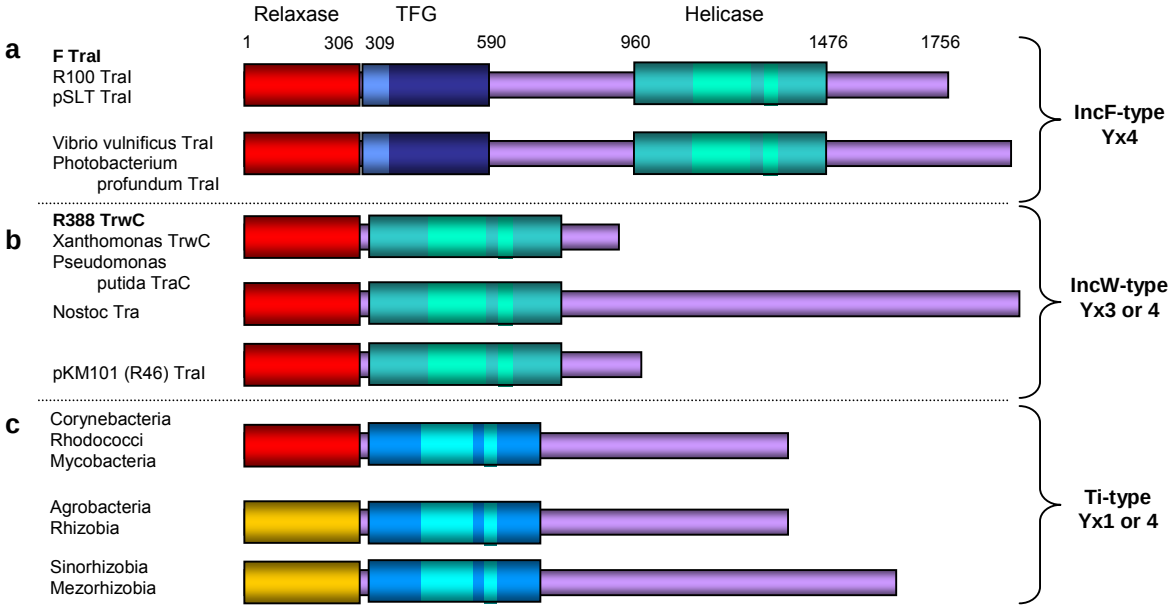
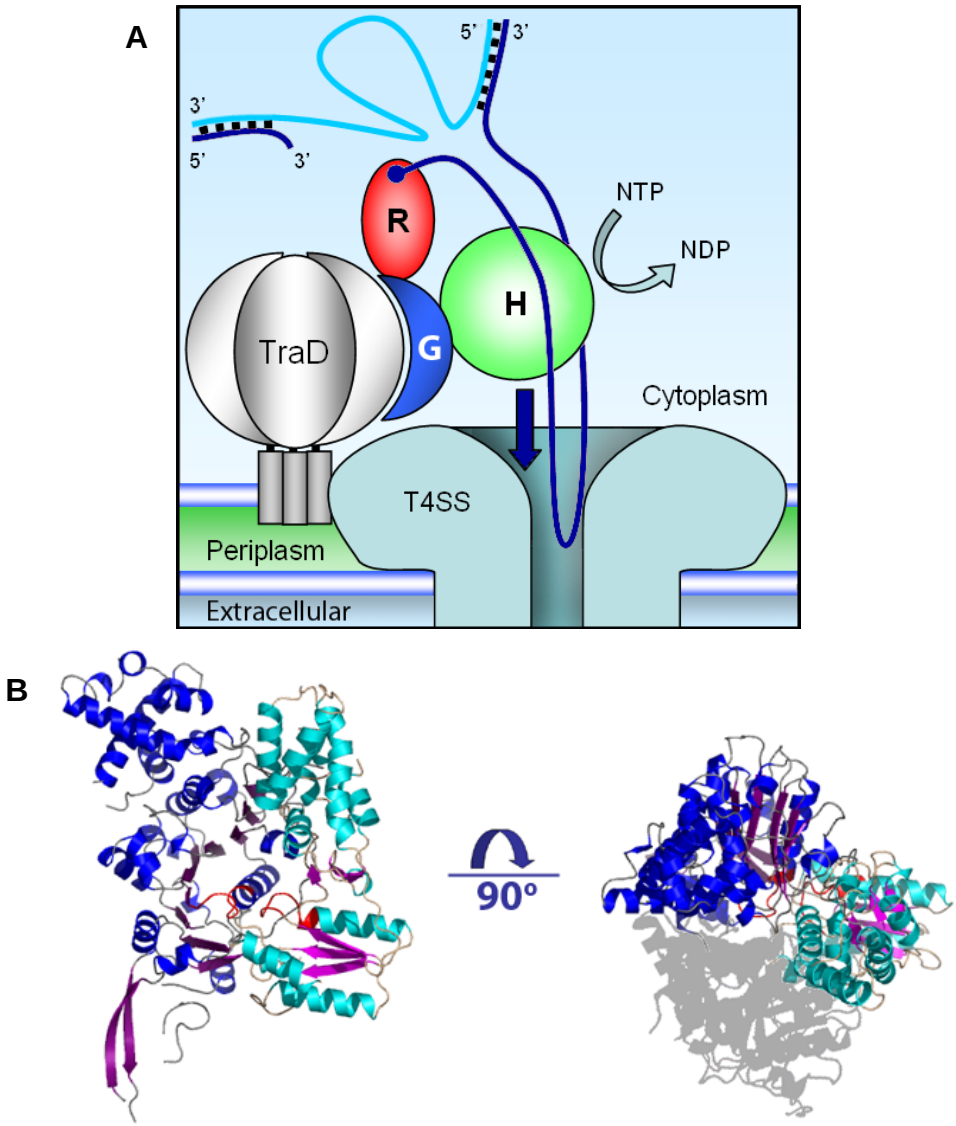


Figure 1-3 Models of Proposed TraI/TraD Interaction



1.8 References

1. Byrd, D. R., Sampson, J. K., Ragonese, H. M., & Matson, S. W. (2002) *J Biol Chem* **277,** 42645-42653.

2. Sampson, J. K. & Matson, S. W. (2004) (University of North Carolina, Unpublished).

3. Grandoso, G., Avila, P., Cayon, A., Hernando, M. A., Llosa, M., & de la Cruz, F. (2000) *J Mol Biol* **295,** 1163-1172.

4. Collignon, P. J. (2002) *Med J Aust* **177,** 325-329.

5. Wilkins, B. M. (2002) *Environ Microbiol* **4,** 495-500.

6. Waters, V. L. (1999) *Front Biosci* **4,** D433-456.

7. Espinal, M. A. (2003) *Tuberculosis (Edinb)* **83,** 44-51.

8. Escalante, P., Ramaswamy, S., Sanabria, H., Soini, H., Pan, X., Valiente-Castillo, O., & Musser, J. M. (1998) *Tuber Lung Dis* **79,** 111-118.

9. Makino, K., Ishii, K., Yasunaga, T., Hattori, M., Yokoyama, K., Yutsudo, C. H., Kubota, Y., Yamaichi, Y., Iida, T., Yamamoto, K.*, et al.* (1998) *DNA Res* **5,** 1-9.

10. Golubov, A., Neubauer, H., Nolting, C., Heesemann, J., & Rakin, A. (2004) *Infect Immun* **72,** 5613-5621.

11. Dyda, F. & Hickman, A. B. (2003) *Structure (Camb)* **11,** 1310-1311.

12. Llosa, M., Gomis-Ruth, F. X., Coll, M., & de la Cruz Fd, F. (2002) *Mol Microbiol* **45,** 1-8.

13. Howard, M. T., Nelson, W. C., & Matson, S. W. (1995) *J Biol Chem* **270,** 28381-28386.

14. Yoshioka, Y., Ohtsubo, H., & Ohtsubo, E. (1987) *J Bacteriol* **169,** 619-623.

15. Rees, C. E. & Wilkins, B. M. (1990) *Mol Microbiol* **4,** 1199-1205.

16. Rees, C. E. & Wilkins, B. M. (1989) *J Bacteriol* **171,** 3152-3157.

17. Inamoto, S., Fukuda, H., Abo, T., & Ohtsubo, E. (1994) *J Biochem (Tokyo)* **116,** 838-844.

18. Nelson, W. C., Howard, M. T., Sherman, J. A., & Matson, S. W. (1995) *J Biol Chem* **270,** 28374-28380.

19. Ragonese, H. M. & Matson, S. W. (2004) (University of North Carolina, Unpublished).

20. Matson, S. W., Sampson, J. K., & Byrd, D. R. (2001) *J Biol Chem* **276,** 2372-2379.

21. Lujan, S. & Redinbo, M. (2004) (University of North Carolina, Unpublished).

22. van Mansfeld, A. D., van Teeffelen, H. A., Baas, P. D., & Jansz, H. S. (1986) *Nucleic Acids Res* **14,** 4229-4238.

23. Datta, S., Larkin, C., & Schildbach, J. F. (2003) *Structure (Camb)* **11,** 1369-1379.

24. Guasch, A., Lucas, M., Moncalian, G., Cabezas, M., Perez-Luque, R., Gomis-Ruth, F. X., de la Cruz, F., & Coll, M. (2003) *Nat Struct Biol* **10,** 1002-1010.

25. Ramirez, U. D., Minasov, G., Focia, P. J., Stroud, R. M., Walter, P., Kuhn, P., & Freymann, D. M. (2002) *J Mol Biol* **320,** 783-799.

26. Doudna, J. A. & Batey, R. T. (2004) *Annu Rev Biochem* **73,** 539-557.

27. Egea, P. F., Shan, S. O., Napetschnig, J., Savage, D. F., Walter, P., & Stroud, R. M. (2004) *Nature* **427,** 215-221.

28. Focia, P. J., Shepotinovskaya, I. V., Seidler, J. A., & Freymann, D. M. (2004) *Science* **303,** 373-377.

29. Gomis-Ruth, F. X., Moncalian, G., de la Cruz, F., & Coll, M. (2002) *J Biol Chem* **277,** 7556-7566.

30. Gomis-Ruth, F. X. & Coll, M. (2001) *Int J Biochem Cell Biol* **33,** 839-843.

31. Lujan, S. A., Guogas, L. M., Ragonese, H., Matson, S. W., & Redinbo, M. R. (2007) *Proceedings of the National Academy of Sciences of the United States of America* **104,** 12282-12287.

2 Inhibition of F Plasmid TraI: Structure, Kinetics, & *in cell* assays

Conjugative transfer of plasmid DNA via close cell-cell junctions is the main route by which antibiotic resistance genes spread between bacterial strains. Relaxases are essential for conjugative transfer and act by cleaving DNA strands and forming covalent phosphotyrosine linkages. Based on data indicating that multi-tyrosine relaxase enzymes can accommodate two phosphotyrosine intermediates within their divalent metal-containing active sites, we hypothesized that bisphosphonates would inhibit relaxase activity and conjugative DNA transfer. We identified bisphosphonates that are nanomolar *in vitro* relaxase inhibitors. Furthermore, we utilized cell-based assays to demonstrate that these compounds are highly effective at preventing DNA transfer and at selectively killing cells harboring conjugative plasmids. Two potent inhibitors, clodronate and etidronate, are already clinically approved to treat bone loss. Thus, the inhibition of conjugative relaxases is a potentially novel antimicrobial approach, one that selectively targets bacteria capable of transferring antibiotic resistance and generating multidrug resistant strains.

2.1 Introduction

Conjugative elements are responsible for the majority of horizontal gene transfers within and between bacterial strains (reviewed (1)), as first described for the *Escherichia coli* F plasmid by Lederberg and Tatum in 1946 (2). Conjugative DNA transfer is also the central mechanism by which antibiotic resistance and virulence factors are propagated in bacterial populations (reviewed (3)). Indeed, it is well established that antibiotic resistance can be rapidly acquired in clinical settings, and that such acquisition is critically dependent on conjugative DNA transfer (reviewed (4)). Small molecule inhibition of conjugation could prove to be a powerful method for curbing the generation and spread of multi-drug resistant strains. Past studies suggested that various antibiotics, polycyclic chemicals, and crude extracts inhibit conjugation at concentrations less than the antibacterial minimum inhibitory concentration (MIC)(5-11); however, most of these effects have been attributed to non-conjugation-specific inhibition of bacterial growth or DNA synthesis(12-15). This study describes a bottom-up approach used to identify the first small molecule inhibitors of conjugative DNA transfer that target an enzyme of the conjugative system.

The DNA relaxase is a central enzyme in each conjugative system (16-18), and thus is a prime target for inhibition. The conjugative relaxase initiates DNA transfer with a site and strand specific single-stranded DNA (ssDNA) nick in the transferred strand (T-strand) at the origin of transfer (*oriT*), forming a covalent

5'-phosphotyrosine intermediate (16, 19-23). The nicked T-strand moves from the donor cell (*plasmid*$^+$) to the recipient cell (*plasmid*$^-$) via an intercellular junction mediated by a Type IV secretion system (reviewed (19, 24, 25)). The relaxase completes DNA transfer by reversing the covalent phosphotyrosine linkage and releasing the T-strand. In the F plasmid, this relaxase is located in the N-terminal domain of a large multifunctional protein, TraI (DNA Helicase I) (22, 23, 26-28). Some conjugative relaxases utilize one active site tyrosine (*e.g.,* IncQ RSF1010 MobA (29), IncP RP4 TraI (30, 31), IncI R64 NikA (32), *Agrobacterium* Ti VirD2 (33), Tn5252 MocA/BmgA (34); where "R" indicates plasmids that propagate antibiotic resistance). F-like relaxases (*e.g.,* IncF R1 and R100 plasmid TraIs (28), IncN R46 and pCU1 TraIs (35), IncW R388 TrwC (36), *Pseudomonas* IncP9 pWW0 TraC (37)) maintain a conserved, bifurcated constellation of two to five tyrosines near their N-termini. The most common arrangement is four tyrosines (Y1-Y4; tyrosines 16, 17, 23, and 24 in F TraI), with pairs Y1/2 and Y3/4 separated by a variable linker region. Crystal structures show all four tyrosines are proximal to a bound metal ion (this study and others (38-41)). Optimal relaxase cleavage, ligation and transfer of ssDNA require the metal ion and two catalytic tyrosines, one from each pair (42). F TraI relaxase shares significant sequence identity with relaxases of many R plasmids (*e.g.*, 98% with R100 TraI); thus, the F plasmid serves as a model system for examining conjugative plasmids and the inhibition of conjugative transfer.

In this study, we first sought to understand the role that the relaxase enzyme plays in the initiation and termination of DNA conjugation, and then sought to use that information to identify potent relaxase-specific inhibitors. Our results establish that the conjugative DNA transfer process can be selectively disrupted by relaxase-targeted compounds, including some clinically-approved drugs. This is a potentially novel antimicrobial approach, one that could be used to purge from microbial populations the bacteria capable of propagating antibiotic resistance genes.

2.2 Results

We determined the 2.4 Å crystal structure of the 300-residue N-terminal relaxase domain of F plasmid TraI (N300) with a tyrosine 16 to phenylalanine mutation (Y1 of F TraI; Y16F) (**Fig. 1a**; **Table 2**). The structure of N300 is similar to those of a 330-residue F TraI fragment (39, 41) (N330) and the relaxase domain of R388 TrwC (38, 40) (**Fig. 5**). Crystallization required a 9-base single-stranded DNA (ssDNA) oligonucleotide consisting of the F *oriT* nick site sequence. In spite of the Y16F mutation, which reduces N300 DNA cleavage 600-fold, we observed electron density for just one DNA base in the active site (**Fig. 1a**). We interpreted this as the *oriT* thymidine immediately upstream of the scissile phosphate (-1 Thy). We also observed density for a bound metal ion, chelated by three conserved histidine sidechains (*soft ligands*, i.e. uncharged and high polarizability). Despite unusual soft chelation, we interpret the bound metal as a divalent magnesium ion (a *hard center*, i.e. low polarizability and a

preference for charged ligands) based on bond lengths, electron density, and octahedral coordination. A survey of magnesium-binding proteins in the Protein Data Bank revealed that the chelation of Mg^{2+} by neutral residues is diagnostic of a site that simultaneously binds to multiple phosphate groups (**Table 2**). Mutation of the metal-chelating residue histidine 159 to glutamic acid (thus reducing the effective charge of the metal site) eliminated relaxase activity (**Fig. 6**). These data indicate that the 2+ charge on the bound metal ion is critical to relaxase function.

We then considered models for conjugative DNA transfer that would require the observed 2+ metal ion and two catalytically-competent tyrosine residues (**Fig. 1b**). In the case of Simple Transfer (green), a single catalytic tyrosine is sufficient for successful intercellular DNA transfer. However, if the free 3'-hydroxyl product of relaxase-mediated DNA cleavage becomes a substrate for Concomitant Plasmid Replication (CPR; purple), analogous to rolling circle replication (43), then two tyrosines would be required to resolve the replicative *oriT* intermediate and release the T-strand to complete transfer. CPR events explain the reported low frequency generation of greater than unit-length conjugative plasmids (44). Although CPR may not be the primary conjugative pathway, plasmids with relaxases capable of resolving CPR intermediates would be expected to have a selective advantage. Single-tyrosine relaxases could achieve resolution of CPR intermediates through relaxase multimerization (45) or cooperation with a second non-relaxase protein (46). The ability to resolve

undesirable replication products would also confer *oriT* specific recombinase activity between plasmids (47, 48) or between tandem *oriT* repeats on the same plasmid (46, 49). However, for our purposes, a key prediction that arose from this model is that multi-tyrosine relaxases are capable of accommodating two phosphotyrosine intermediates simultaneously within their active sites. The need to handle dual phosphotyrosine intermediates would explain the key features of the TraI relaxase outlined above: the presence of two catalytically-competent tyrosines and the requirement for a metal ion with an obligate 2+ charge.

We examined the catalytic role of the relaxase active site divalent metal ion, which is coordinated by three conserved histidine residues and has been interpreted as either magnesium or zinc (40, 41). While zinc coordination by histidines is common, the exclusive chelation of a hard magnesium ion by soft imidazole side chains is unusual. However, magnesium may be essential to relaxase function, as the activity of several relaxases are enhanced by Mg^{2+} (29, 50). As stated in the main body of this study, a survey of magnesium-binding proteins in the Protein Data Bank (51) revealed that the chelation of Mg^{2+} by neutral residues is diagnostic of polyphosphate binding (**Table 2**). We tested whether the 2+ charge on the metal ion bound to the F TraI relaxase was required for enzyme function by mutating one of the three coordinating histidine residues (H159) either to a neutral glutamine residue (H159Q) or to an acidic glutamic acid residue (H159E). The nearly isostructural H159Q mutation had no effect on enzyme function, generating cleavage and crossover activity

approximately 95% of wild type TraI. The H159E mutation, however, which would reduce the effective charge on the bound metal ion to 1+, essentially eliminated enzyme function, generating a 12-fold decrease in cleavage activity, no measurable crossover activity, and a three order-of-magnitude decrease in conjugative transfer activity. Thus, a 2+ charge on the bound metal ion is essential for the catalytic activity of the TraI relaxase, particularly the generation of DNA crossovers and successful conjugative DNA transfer.

Based on the prediction that multi-tyrosine relaxases employ two simultaneous phosphotyrosine intermediates coordinated to one magnesium ion, we hypothesized that simple bisphosphonates would bind the magnesium center and thus inhibit F TraI relaxase activity and conjugative transfer. To test this hypothesis, a kinetic assay using fluorophore-labeled *oriT* ssDNA for cleavage by the F TraI relaxase domain (TraI N300) was developed to complement existing radiolabel-based techniques (52, 53). Imidodiphosphate (PNP), a simple and relatively stable bisphosphonate, was the first compound examined in this assay (**Fig. 2a**). We found that PNP is a nanomolar inhibitor of TraI relaxase activity *in vitro*. Analysis of cleavage velocity curves revealed that PNP is a mixed-type (specifically, noncompetitive) inhibitor of TraI, with apparent competitive ($K_{ic,app}$) and uncompetitive ($K_{iu,app}$) inhibition constants of 2.0-2.4 nM and 2.7-3.5 nM, respectively. Thus, a simple bisphosphonate serves as a potent inhibitor of a multi-tyrosine relaxase activity *in vitro*.

To determine whether PNP derives its inhibitory power by binding to the TraI relaxase active site, TraI N300 Y16F crystals were soaked with PNP and the x-ray structure determined to 3.0 Å resolution (N300+PNP; **Table 1**). As in the N300 structure, three of the six magnesium octahedral coordination positions are filled by histidine side chains, the fourth by the 3'-hydroxyl of the scissile thymidine, and the fifth is occluded by the Y16F side chain (**Fig. 2b**). Unlike the N300 structure, a 5σ simulated-annealing omit electron density peak appears in the sixth coordination position (**Fig. 8**), indicating the binding of a single PNP phosphate group within 3.7 Å of the magnesium ion. The second PNP phosphate was not observed, due either to disorder or, more likely, to hydrolysis by a water molecule activated by the adjacent 2+ metal (54-57). The N300+PNP structure supports the conclusion that PNP inhibits TraI by binding to the relaxase catalytic site.

The N300+PNP structure revealed a novel phosphate binding site, which allowed us to model the second phosphotyrosine intermediate in the relaxase active site. By rotating the first alpha helix (αA; lower right corner **Fig. 1a**) to match the orientation observed in the R388 TraI homologue, TrwC (40), and extending helicity through tyrosine 24 (Y4 of F TraI), tyrosine 23 (Y3 of F TraI) reorients such that its side chain hydroxyl overlaps with the N300+PNP phosphate position (**Fig. 2c**). In this orientation, tyrosine 24 also makes an aromatic stacking interaction with the side chain of tryptophan 278 (W278), a residue that is conserved in relaxases that have Y3/4 pairs. This model explains

how two phosphotyrosines, one at Y16 and one at Y23, can be accommodated within the active site, and in the process fulfill the octahedral coordination geometry of a bound magnesium ion.

A variety of compounds were then examined for their ability to inhibit the F TraI relaxase *in vitro.* A coarse screen (200 nM concentrations and pH 7.4) of eleven bisphosphonates and three negative controls (sodium chloride, dibasic potassium phosphate, and ampicillin, none of which exhibited relaxase inhibition) yielded five additional inhibitors: methylenediphosphonic acid (PCP); iminobis(methylphosphonic acid) (PCNCP); etidronic acid (ETIDRO); clodronic acid (CLODRO); and 1,2-bis(dimethoxyphosphoryl)benzene (PBENP) (**Fig. 3**). Results from this screen reveal that effective inhibitors have two phosphonate moieties separated by three or fewer atoms and have no additional negative charge at pH 7.4. Four clinically-approved bisphosphonates were tested; these drugs are used to treat bone loss by inhibiting farnesyl diphosphate synthase (reviewed (58)). The simplest, ETIDRO and CLODRO, inhibited the TraI relaxase, while pamidronic acid and neridronic acid (PAMDRO, NERDRO, respectively), which have an alkyl-amine side chain, did not. Two other inhibitors identified, PCP and PNP, have been used as radioisotope carriers in humans (59, 60). Pyrophosphate was not examined due to its rapid hydrolysis in aqueous solution. The simplest inhibitors, PCP, ETIDRO and CLODRO, were then characterized further using a kinetic assay and exhibited purely competitive inhibition, with $K_{ic,app}$ values ranging from 3-145 nM (**Fig. 9**). Taken together with

the PNP results, these data validate the prediction that F-like conjugative relaxases can accommodate two phosphotyrosine intermediates simultaneously within their active sites. Significantly, these data also establish that bisphosphonates (including clinically-approved compounds) potently inhibit the *in vitro* relaxase activity of F TraI with K_i values in the nanomolar range.

We next asked whether PNP could impact conjugative DNA transfer between living bacterial cells. F^+ *E. coli* were mated with F^- *E. coli* in dilute media and in the presence of increasing concentrations of PNP. The resulting mixture was applied to agar plates with antibiotic selection for transconjugants (newly formed F^+ cells). Colony counts revealed that PNP inhibited DNA transfer with a half-effective concentration (EC_{50}) of approximately 10 μM, the lowest concentration tested in this assay (**Fig. 4a**). We also found that PNP selectively kills F^+ donor cells with an EC_{50} of <10 μM (compared to low millimolar EC_{50} against F^- recipient cells). This suggests that the TraI relaxase sensitizes F^+ cells to a toxic effect of PNP. Minimal cell growth was observed in control mating mixtures, so decreases in donor cell count relative to controls are attributed primarily to cell death rather than to a lack of cell growth. Thus, inhibitor-dependant decreases represent bacteriacidal rather than bacteriastatic effects. Strains containing an F plasmid lacking the TraI gene (F^+/$TraI^-$) or containing an F plasmid with the four active site tyrosines mutated to phenylalanine (F^+/TraI 4Y-F) behaved like F^- cells in this assay, in that they were resistant to the lethal effects of PNP. The impact on DNA donor cell survival depends on the presence

of a catalytically active TraI relaxase. Thus, the simple bisphosphonate PNP enters living bacteria, inhibits conjugative DNA transfer, and selectively kills cells in an active relaxase-dependent manner.

A fluorescence-based 96-well assay was then employed to examine further these effects and to screen additional compounds for the ability to impact cell survival and DNA transfer in living bacterial cells. This higher-throughput cell enumeration assay utilized an oxygen-quenched fluorophore imbedded within a hydrophobic gel; the concentration of live (oxygen consuming) cells is proportional to fluorescence (61). The six compounds effective at inhibiting TraI relaxase activity *in vitro* (PNP, PCP, PCNCP, PBENP, CLODRO, ETIDRO; **Fig. 3**), along with two controls (PAMDRO and K_2HPO_4 [PO4]) were examined for their impacts on F^+ and F^- cell survival and on DNA transfer (**Fig. 4b, 4c**). PAMDRO and PO4, which had no effect on TraI relaxase activity *in vitro*, showed little effect on cell survival and DNA transfer in these cell-based assays, exhibiting EC_{50} values greater than the highest concentrations tested (10 mM and 100 mM, respectively) and no selectivity for F^+ over F^- cells. In contrast, all six potent *in vitro* TraI relaxase inhibitors were also effective in living *E. coli* cells. EC_{50} values for inhibiting F^+ donor cell survival ranged from 10 μM (ETIDRO) to 16 nM (PCNCP), and for inhibiting conjugative DNA transfer from 31 μM (PNP) to 110 nM (CLODRO). These compounds have little effect on F^- recipient cells, with EC_{50} values ranging from 0.34 mM (ETIDRO) to >100 mM (PBENP), which represents 30- to 10^6-fold selectivity for F^+ cells, respectively. In general, PNP

and PBENP were more effective at inhibiting F^+ cell survival, ETIDRO and CLODRO were more effective at inhibiting DNA transfer, and PCP and PCNCP were effective against both F^+ cell survival and DNA transfer (**Fig. 4b, 4c**; **Fig. 3**). The ranges for EC_{50} values (**Fig. 4b**) were derived considering the error bars present in the survival curves (**Fig. 4c**). Thus, for the most extreme case of ETIDRO, the median EC_{50} for transfer inhibition is 330 nM with a range of 1.1 μM to <10 pM (**Fig. 10**). Taken together, however, these data establish that relaxases can be inhibited with nanomolar affinity within living bacterial cells and that this inhibition both limits DNA transfer and selectively kills microbes harboring conjugative plasmids.

2.3 Discussion

This study outlines a potentially novel antimicrobial paradigm that specifically targets the DNA relaxase enzyme required to initiate and terminate the process of bacterial conjugation. The compounds identified could be used along or in combination with existing antibiotics to treat recalcitrant bacterial infections. Other antibiotics and natural extracts have been reported to disrupt conjugative DNA transfer and the presence of plasmids within actively dividing bacterial cells(5-15). In each case, however, the macromolecular target of those compounds was not understood and their mechanism of action has not been determined. We took a “bottom up” approach targeted DNA conjugation by considering first the mechanism and role of a single enzyme (the DNA relaxase) in this process, and then identifying inhibitors to test a specific mechanistic hypothesis (that two phosphotyrosine intermediates can be accommodated in the

relaxase active site). While our data was collected on the relaxase from the well-established F plasmid that was first identified in 1946(2), the F conjugative machinery shares up to 99% sequence identify with R plasmids that transfer antibiotic resistance in the wild. Thus, our approach is likely to be effective against a range of plasmids involved in propagating a range of resistance genes and virulence factors, many of which play an important role in clinical infections (62-64).

In addition to inhibiting DNA transfer, we show that simple bisphosphonates selectively purge populations of bacteria containing a conjugative plasmid with an active relaxase enzyme. However, because TraI is F plasmid encoded, not an essential *E. coli* enzyme, and was not expected to play a significant role in isolated F^+ cells, this relaxase-dependent cell lethality was a surprise. Several mechanisms could be envisioned to explain this observation, and future studies will be required to distinguish between them. One possibility is that relaxases engage in cycles of plasmid DNA cleavage and religation that are uncoupled from mating and conjugative DNA transfer, and that the disruption of that process results in a competitive disadvantage relative to cells without conjugative plasmids. Indeed, inspection of **Fig. 4C** reveals that a direct competition appears to exist between plasmid propagation and donor cell survival (note particularly PNP, PCNCP, PBENP, CLODRO). F+ donor cell survival is enhanced at higher bisphosphonate concentrations to the detriment of plasmid propagation via conjugative transfer. These observations suggest a

classic “zero-sum game” in which toxic relaxase-specific bisphosphonate inhibitors pit the interests of endosymbiont plasmids against those of their bacterial hosts.

We show that the clinically-approved bisphosphonates etidronate (Didronel®) and clodronate (Bonefos®), but not other bisphosphonate therapeutics, are potently effective at killing F^+ cells and preventing conjugative DNA transfer. These particular compounds could also be combined with existing antibiotics to create potent antimicrobial cocktails. Etidronate and clodronate exhibit low absorption (65, 66) and can be administered at high oral doses (**Table 3**). Extrapolating from our results, approved doses of etidronate and clodronate would be expected kill >90% of *plasmid*$^+$ cells and to stop >80% of conjugative transfer within the gastrointestinal track. Such results are relatively mild, given the large bacterial populations present in the gastrointestinal tract or at wound sites, but may be enough shift the balance toward success in a variety of recalcitrant clinical infections, especially given the prevalence of conjugative plasmids within multi-drug resistant bacterial strains. The treatment of skin infections, primary sites of nosocomial antibiotic resistance transfer, using the topical applications of bisphosphonates may also be effective. In summary, this study establishes conjugative relaxases as a unique antimicrobial target. Our results suggest that approved therapeutics could have an immediate impact, alone or in combination with existing antibiotics, in the prevention of resistance

propagation during clinical treatment of bacterial infections, and in extending the lifetime of our antibiotic arsenal.

2.4 Materials and Methods

2.4.1 Protein Expression and Purification

An amino-terminal 300 residue F plasmid TraI construct, bearing a tyrosine to phenylalanine mutation at position 16 (N300 Y16F), was cloned into IMPACT® vector pTYB2 (New England Biolabs) for expression as a C-terminal intein-chitin-binding-domain (CBD) fusion. Protein was expressed in either *E. coli* BL21 (DE3)/pLysS or HMS174 (DE3)/pLysS and was purified as per the standard IMPACT® protocol. Briefly, cellular extracts were prepared and incubated with 1 mL of Chitin Resin (New England Biolabs) per liter of cell culture. The resin was washed and incubated with 50 mM dithiothreitol (DTT) overnight to cleave the relaxase from its CBD tag. The DTT laden eluent was extensively dialyzed in 20 mM NaCl and 20 mM Tris-HCl (pH 7.5). The resulting N300 Y16F was >95% pure by SDS-PAGE, and was concentrated to 3 mg/mL for crystallization in 50 mM NaCl, 10% glycerol, and 10 mM Tris-HCl (pH 7.5) prior to flash-freezing in liquid nitrogen for storage at -80 °C. Protein for functional and kinetic assays was concentrated to 42.3 μM in 150 mM NaCl, 50% glycerol, and 10 mM Tris-HCl (pH 7.5) for long-term storage at -80 °C.

2.4.2 Oligonucleotides

A 9-base single-stranded DNA oligonucleotide (9mer) derived from the F plasmid *oriT* (5'-GGT GT^G GTG-3', where ^ is the scissile phosphate) was synthesized for crystallization at the UNC Lineberger Comprehensive Cancer

Center Nucleic Acids Core Facility. Labeled oligonucleotides for fluorescence fluorescent kinetic assays were synthesized by Integrated DNA Technologies (IDT). For the initial kinetic assay the substrate oligonucleotide was a 5'-biotin (bio) labeled 29mer ("b29"; 5'-bio-TTT GCG TGG GGT GT^G GTG CTT TTG GGT GG-3'). The complementary fluorescent probe oligonucleotide was a 5'-6-carboxyfluorescein (56FAM™) labeled 15mer ("downF"; 5'-56FAM-CC ACC CAA AAG CAC C-3'). The b29 substrate molecule was derived from the F plasmid *oriT*. The downF probe is complementary to the downstream portion of b29. The melting temperature was 50.8° C for downF versus b29 (15 base-pairs) , as calculated with IDT OligoAnalyzer™ 3.0 with default parameters. For later kinetic assays, the substrate oligonucleotide was a 5'-biotin (bio) labeled and 3'-56FAM labeled 27mer ("Bio19'8FAM"; 5'-bio-CT TGT TTT TCG TGG GGT GT^G GTG CTT T -3'). The Bio19'8FAM substrate molecule has a larger portion of the F plasmid *oriT* than does the b29 oligonucleotide. Oligonucleotides for site directed mutagenesis were synthesized by Integrated DNA Technologies.

2.4.3 Crystallization and Structure Determination

N300 Y16F crystals grew in a DNA-dependent manner in 75 mM sodium nitrate, 14% w/v PEG 3350, 10 mM spermine, and 110 μM 9mer. These rods were cryoprotected via a two-second dip in 150 mM sodium nitrate, 35% w/v PEG 3350, and 10 mM spermine and flash cooled in liquid nitrogen for storage and transport. Crystals employed for the PNP-bound structure were soaked for 24 hours in 200 mM ammonium nitrate, 40% w/v PEG 3350, and 1 mM N,N-

imidobisphosphonate (PNP) and flash cooled. Rods 200 × 30 × 20 μm in size were generated by hanging drop vapor diffusion (>35 days of growth) and diffracted to between 2.9 Å and 3.4 Å in-house. Data sets were collected at the Advanced Photon Source (APS) at Argonne National Laboratoy (ANL), at Southeast Regional Collaborative Access Team (SER-CAT) Sector 22 Insertion Device Beamline (22-ID) and the General Medicine and Cancer Institutes Collaborative Access Team (GM/CA-CAT) Sector 23 Insertion Device Beamline (23-ID_{in}; for the PNP complex). Crystals were of space group of $P2_12_12_1$ and contained two protein monomers in the asymmetric unit (**Table 1**). X-ray diffraction data were indexed and scaled with the HKL2000 or MOSFLM (CCP4) (67). Initial phases were determined by molecular replacement in Molrep (CCP4) (67) with the apo-TraI structure (Protein Data Bank accession 1P4D (41)) as a search model. Model adjustment was completed with O (68) and σ_a-weighted electron density maps (69), and structures were refined using torsion angle dynamics and the maximum likelihood target as implemented in CNS (70). Structure figures were constructed in PyMol v0.98 (71).

2.4.4 Functional Assays

Both wild-type and mutant TraI proteins (either full length protein or TraI-N300) were examined in oligonucleotide cleavage (DNA nicking), strand transfer (DNA religation) and liquid mating (DNA transfer) assays. The oligonucleotide cleavage reaction mixture (10 μl) contained 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 6 mM $MgCl_2$, 20% glycerol, 1 pmol 5'-end labeled 22-mer, and 1 pmol TraI

(unless otherwise stated). Reactions were assembled at room temperature and incubated at 37°C for 20 minutes. Reactions were stopped by the addition of SDS to 0.2%, and incubation was continued at 37°C for 10 minutes. Ten µl 85% formamide, 50 mM EDTA, 0.1% dyes were added to the reaction, the products were denatured at 100°C for 3 minutes and analyzed on a 16% polyacrylamide, 8 M urea denaturing gel. The gels were electrophoresed at 25 watts in 1xTBE (90 mM Tris-borate and 2 mM EDTA) and visualized using a PhosphorImager (Molecular Dynamics). Markers were prepared as described previously (72). Strand transfer reactions were performed in a manner similar to the oligonucleotide cleavage assay except after the 20-minute incubation, 1 pmol of a second unlabeled oligonucleotide of differing length containing the F plasmid *nic* site was added to the reaction and incubation was continued at 37°C for 1 hour. The reaction was stopped and analyzed using the procedure described above. Liquid mating assays were performed as previously described (26) except HMS174 cells were utilized instead of HMS174 (DE3) to reduce the constitutive expression of the complementing protein. Briefly, cells containing pOX38TΔTraI and the appropriate complementing plasmid were grown overnight in the presence of appropriate antibiotics. Overnight cultures were used to inoculate cultures that were grown at 37^{o}C to mid-log phase (2-3 hours) in the absence of antibiotics. Donor cells were mixed (1:10) with recipient cells, incubated at 37^{o}C and then plated to select for transconjugants and counterselect for donors and recipients. Site-directed mutations in the *traI* gene were created using mutagenic primers and the site-directed mutagenesis

protocol supplied by Stratagene. pTYB2-traIN300 served as the template for PCR. The resulting clones were sequenced to confirm the presence of the engineered mutations and the absence of unintended mutations. A unique 700 bp *Nde*I-*Stu*I fragment of *traI* containing the engineered mutations was removed from pTYB2-*traI*N300 and ligated into the full length *traI* gene in pET11c-traI that had been digested at unique *Nde*I and *Stu*I sites to create the mutant pET11c-traI derivatives that were at utilized in genetic complementation assays.

2.4.5 Kinetic Assays

2.4.5.1 Kinetic Assay Formulations

Reaction Buffer: 6.42 mM $MgCl_2$, 20.5% glycerol, 153.9 mM NaCl, and 51.3 mM Tris-HCl pH 7.5. *Streptavidin Wash Buffer* (Tris buffered saline, TBS/Tween): and 150 mM NaCl, 25 mM Tris-HCl pH 7.5, and 0.05% Tween-20. *Stopping Buffer*: 1.2% sodium-dodecyl sulfate (SDS), and 300 mM EDTA pH 10. TE Buffer: 1 mM EDTA, and 10 mM Tris-HCl pH 7.4. *Fluorescence Buffer*: 80% glycerol, 200 mM Tris-HCl pH 8.0. *Short-term N300 stock* (for storage at -20 °C): 50% Reaction Buffer, 49.8% glycerol, 0.2% long-term protein solution (84.6 nM final N300 concentration). All 5x stocks except Probe stocks were diluted in Reaction Buffer. *5x Enzyme Stock*: 8.4 mL short-term N300 stock diluted to 2.02 nM N300. *5x Inhibitor Stocks*: imidodiphosphate (PNP) at 0-50 nM in Reaction Buffer. *5x Substrate Stock*: b29 oligonucleotide was diluted to 19.6-158.5 mM each, by 3-fold serial dilutions. *5x Probe Stocks*: 23.5-190.2 mM (1.2-fold molar excess) downF oligonucleotide in TE Buffer. All reactions and procedures

involving downF were assembled and performed in low-light conditions. Solutions and microtiter plates containing downF were kept in foil-lined containers at all times to prevent photobleaching. In *Later* kinetic assays, different substrate dilution schemes were used, Fluorescence Buffer was abandoned in favor of Streptavidin Wash Buffer, and the Stopping Buffer SDS concentration was reduced to 0.2%.

2.4.5.2 Fluorescent Kinetic Assays

Two *oriT* derived oligonucleotides, b29 and Bio19'8FAM, were designed for binding and cleavage by TraI based on past studies (73) in two phases. The *Initial* phase includes the initial uninhibited characterization of N300 oligonucleotide cleavage, all inhibition studies with PNP, and a blind rough screen of all listed inhibitor candidates (**Fig. 3**). The *Later* phase used improved oligonucleotide design and included the inhibitor characterizations of PCP, ETIDRO, and CLODRO. The overall method is similar to those described previously for the study of TraI and R388 TrwC (27). Reactions were assembled from 16 µL of 5x Substrate Stock, 16 µL of 5x Inhibitor Stock, and 32 µL of Reaction Buffer. 80 µL reactions were initiated with 16 µL of 5x Enzyme Stock and raised to 37 °C. 10 µL samples were removed at eight time points and stopped. In *Initial* assays, reactions were stopped with 10 µL of Stopping Buffer at room temperature and placed on a 100 °C heat block for one minute, spiked with 2 µL of 5x Probe Stock while still hot, and incubated at 37 °C for 10 minutes (see *Kinetic Assay Formulations* for probe handling). In *Later* assays, reactions

were stopped with 2 µL of Stopping Buffer at 80 °C. Samples were diluted to 65 µL with Streptavidin Wash Buffer and transferred to streptavidin-coated microtiter plates (Thermo Electron Biobind™ Assembly white 96-well plates for *Initial* assays; Reacti-Bind™ High Binding Capacity black 384-well plates for *Later* assays). Plates were incubated at 37 °C for 45 minutes (*Initial* assays) or one hour (*Later* assays). Plates were washed with at least 5-fold excess Streptavidin Wash Buffer (in an inverted position to prevent excess biotinylated species from transferring between wells; in *Later* assays this step occurred after reading the plate to determine total oligonucleotide content). Washed wells were then filled with either 65 µL of Fluorescence Buffer (for optimum 6-FAM™ fluorescence; *Initial* assays) or with one well volume of Streptavidin Wash Buffer (*Later* assays). Plates were read in a BMG Labtech FLUOstar Optima (*Initial* assays) or Pherastar (*Later* assays) with a 490 nm excitation and 520 nm emission filters (10 nm bandpass) with the gain optimized for maximum signal.

2.4.5.3 Kinetic Data Processing

All timecourses were fitted by nonlinear regression with un-weighted least-squares methods using SigmaPlot 8.0 (also used for graph construction). Timecourses were manually culled for very high signal (incomplete washing) or very low signal (physical damage to streptavidin coating). Timecourses were scaled to reflect fractional cleavage and replicate timepoints were averaged (n = 1-8; mean n = 3.52). The averaged curves were fitted to a simple, three-

parameter exponential decay equation (Equation 1), and converted into units of concentration.

Equation 1:

$$signal = y_0 + ae^{-bt}$$

where *t* is time in seconds, y_0+*a* is the initial signal, *b* is a constant that determines the shape of the curve. Integration of the overall rate equation for fitting of timecourses was prohibitive. The derivative of the decay equation at zero time was taken as the negative of the initial reaction velocity (v_0; equals *ab* from Equation 1). Michaelis/Menton curves (v_0 versus substrate concentration) were constructed for each inhibitor concentration.

2.4.5.4 Calculation of Kinetic Constants

Competitive (K_{ic}) and uncompetitive (K_{iu}) inhibition constants were estimated from V_{max}^{app} and K_m^{app} values obtained with one of two methods: *first*, velocity curves were fitted to the simple modern Michaelis/Menton equation (Equation 2) by nonlinear regression with un-weighted least-squares method in SigmaPlot 8.0; *second*, the Cornish-Bowden/Eisenthal scale-free direct linear plot method was applied as implemented in the Exploratory Enzyme Kinetics Macro for SigmaPlot 8.0 (74-78). These methods produced virtually identical results, so in *Later* assays, only the former method was used.

Equation 2:

$$v_0 = \frac{V_{\max}^{app} S}{K_m^{app} + S}$$

where K_m^{app} is the apparent Michaelis constant and V_{max}^{app} is the maximum reaction velocity. For PNP, competitive (K_{ic}) and uncompetitive (K_{iu}) inhibition constants were estimated from extrapolation of the negative of ordinate intercepts of K_m^{app}/V_{max}^{app}-versus inhibitor concentration and $1/V_{max}^{app}$-versus-inhibitor concentration, respectively. Definitions for V_{max}^{app} and K_m^{app} (Equations 4-5) were obtained by comparing the equation for mixed inhibition (Equation 3) with the Michaelis/Menton equation (Equation 1) and setting K_{ic} and K_{iu} individually to infinity.

Equation 3:

$$v = \frac{V_{\max} S}{K_m(1 + I/K_{ic}) + S(1 + I/K_{iu})}$$

where *v* is reaction velocity, *S* is substrate concentration, and *I* is inhibitor concentration.

Equation 4, 5:

$$V_{\max}^{app} = \frac{V_{\max}}{1 + I/K_{iu}}, \; K_m^{app} = \frac{V_{\max}/K_m}{1 + I/K_{ic}}$$

where V_{max} is the uninhibited maximum velocity and K_m is the uninhibited Michaelis constant. For inhibitors other than PNP, no uncompetitive inhibition was observed. K_{ic} for these inhibitors was estimated by fitting inhibited velocity curves to Equation 3 with V_{max} and K_m set as determined in uninhibited assays and K_{iu} set to zero.

2.4.6 Mating and Cell Toxicity Assays

2.4.6.1 Mating, selection on solid substrate

Mating assays were performed as described (26). All cell strains are *E. coli* variants. Briefly, the donor strain (JS10) contains pOX38T, a tetracycline resistant and conjugation capable mini-F-plasmid, while the recipient strain (JS4) is F- and streptomycin resistant. Donor and recipient strains from saturated, antibiotic selected overnight cultures were diluted 1:50 and then grown at 37 ºC in the absence of selection to an OD_{600} of approximately 0.6. Donors and recipients were then mixed at a ratio of 1:9 in the presence of the desired concentrations of PNP (*e.g.,* 10, 100 µM, 1, 10 mM) and incubated at 37ºC for five minutes. Mating mixtures were diluted tenfold into broth containing the desired PNP concentration and incubated for an additional 30 minutes. Following mating, the mixtures were vortexed to disrupt mating pairs and serially diluted into 0.9% sterile saline. Dilutions predicted to result in 300-500 colonies on positive control plates were plated on agar plates containing both streptomycin and tetracycline in order to select for transconjugants. Eight replicate samples of each final dilution were plated and the colony counts culled for outliers at 95% confidence and then averaged.

2.4.6.2 Cell Toxicity, selection on solid substrate

These assays proceed as per the Mating Assays above, save that strains are not mixed and thus no mating occurs. In addition to JS4 and JS10 strains, these assays used a strain with a pOX38T containing an inactive TraI gene

(JS11) and a derivative of this strain that was co-transformed with a complementing pET11d plasmid containing a TraI mutant (JS11quad; all four TraI active tyrosines mutated to phenylalanine). JS11 cells have tetracycline resistance. JS11quad cells are tetracycline and ampicillin resistant.

2.4.6.3 Fluorescent Mating & Toxicity, selection in liquid media

Growth, mating and preparation were performed as described above. Mating was performed in the presence of various inhibitors from low picomolar to high millimolar concentrations. Substances shown to be poor *in vitro* inhibitors of F TraI N300 (potassium phosphate and pamidronate) were used as negative controls. Selections for donors, recipients and transconjugants were performed in Oxygen Biosensor™ 96-well round bottom plates (BD Biosciences). Cell concentrations were quantified by fluorescent doubling time analysis as previously described (61). Fluorescence was quantified on a Pherastar fluorescent plate reader (BMG).

2.5 Figure Legends

Figure 1. F Tral N300 Y16F, Bound to the Scissile Thymidine, and a Two-Path Model of F-like Bacterial Conjugation.

A. N300 Y16F active site without PNP (2.4 Å resolution, orange sticks and ribbons). A metal ion (blue sphere) is chelated by three histidines and the -1 Thy nucleotide 3'-hydroxyl. Phenylalanine 16 (Y16F) occludes a fifth octahedral coordination site. Y16F is at the C-terminal of the first alpha helix (αA).

B. Conjugation is initiated by *oriT* on the T-strand (red strand) cleavage by the first tyrosine, forming a covalent phosphotyrosine intermediate (red circle). Transfer with Concomitant Plasmid Replication diverges from Simple Transfer when the free 3'-hydroxyl, left by the initial cleavage, becomes a substrate for replication (blue strand). Replication creates a new *oriT*, necessitating resolution by a second cleavage event and phosphotyrosine formation (purple circle) before ligation of the original T-strand. This results in a dual-phosphotyrosine intermediate.

Figure 2. Relaxase Inhibition by PNP.

A. N300 *oriT* ssDNA cleavage velocity (v_0) inhibited by 0, 1, and 10 nM imidodiphosphate (PNP; red, orange, and yellow, respectively; error bars represent compounded standard errors from timecourse parameter estimates; $n>3$). Competitive/uncompetitive inhibition constants (K_{ic}/K_{iu}) and uninhibited Michaelis constant/maximum velocity (K_m/V_{max}) are from

nonlinear regression and Cornish-Bowden/Eisenthal direct linear plot analyses. Velocities for 100 and 10,000 nM PNP (dark and light blue crosses, respectively), estimated from low signal, were excluded from calculations.

B. N300 Y16F active sites with PNP (3.0 Å resolution; red) and without (orange). A metal ion (blue sphere) is chelated by three histidines, the -1 Thy 3'-hydroxyl, and one PNP moiety (purple). Y16F occludes the sixth coordination site. Residues 236-263 and 266 were disordered (small orange spheres).

C. Dual phosphotyrosine intermediate conformation model (red) constructed from the N300+PNP structure. Helix αA was rotated ~120° about the helical axis to match that of TrwC structure 1OMH, helicity extended through αA', and kinked about histidine 146. Important side chains (red sticks), PNP (purple sticks), a hypothetical second phosphate (purple circle), the bound metal (blue sphere), and the scissile thymidine (from 1OMH; blue) are shown over the TraI active site cleft molecular surface.

Figure 3. Bisphosphonates Examined for Relaxase Inhibition. Chemicals examined for inhibition of TraI activity and F conjugation and for toxicity versus *E. coli* strains. Boxed chemicals were potent *in vitro* TraI inhibitors. *In cell* testing showed that PNP / PBENP were most effective at decreasing F^+ population, CLODRO / ETIDRO at decreasing transconjugant population, and PCP / PCNCP were effective at both.

Figure 4. Effects of F TraI inhibitors on *E. coli* survival and conjugation. F^- (orange lines/bars/symbols), F^+ (red), $TraI^-$ (blue), TraI 4Y-F cells (green), and transconjugant (DNA transfer; yellow).

A. Colony counts after imidodiphosphate incubation (normalized versus uninhibited controls; averaged). Error bars represent standard deviations of the mean (SDOM; $n \geq 5$).
B. EC_{50}s from relative cell counts. Colored bars indicate averaged curve EC_{50}s. Error bars represent ±1 SD envelope EC_{50}s (**Figure S6**). Potent *in vitro* inhibitors are located above a strong black line, negative controls below. Selectivities are EC_{50} ratios for given inhibitors and strains.
C. Relative cell counts with median EC_{50}s (drop arrows; colored by strain). Error bars represent SDOM (n≥3).

Figure 5. Superimposed TraI and TrwC active sites.

A. A superimposition of the N300 Y16F structure and TrwC structure 1OMH (4) (blue backbone cartoon and sidechain sticks) highlights a conformational shift about alpha helix A. This shift rearranges the catalytic tyrosines. The 3'-hydroxyl position of the scissile thymidine (Thy25 in 1OMH) from all DNA-bound TraI and TrwC structures (not all shown) varies by less than 0.5 Å, though the base moiety position varies greatly (face-to-face ring stacked with Y16F in N300 Y16F).

B. A superimposition with TraI N330 2A0I (5) (yellow) highlights a loss of order C-terminal to alpha helix A, including active site tyrosines 23/24, upon extended ssDNA binding.

Figure 6. TraI 300 wild type vs. TraI 300 H159E Cleavage Activity.
Cleavage activity, percentage of *oriT* oligonucleotide cleaved by TraI constructs, as measured from radiological intensities of cleaved and uncleaved ssDNA bands on an SDS gel.

Figure 7. Stereo ribbon cartoon of the N300 active with -1 Thy density.
The N300 Y16F active site of the 2.4 Å structure shows a metal ion (large blue sphere) chelated by a histidine triad and the 3'-hydroxyl of the thymidine base 5' to the scissile phosphate (-1 Thy). The sidechain of phenylalanine 16 (Y16F; replaces wild type tyrosine 16) occludes a fifth octahedral metal coordination site. Backbone atoms are shown as a ribbon cartoon (orange) and important side chains and -1 Thy are shown as sticks (orange). The -1 Thy is surrounded by 3 Å SA-omit density (green cage). Residues between 235 and 264 and between 265 and 267 were disordered (small orange spheres).

Figure 8. The N300 Y16F active site of the 3.0 Å structure with PNP.
One imidodiphosphate (PNP) moiety chelates a metal ion (large blue sphere) along with the histidine triad, and the 3'-hydroxyl of the scissile thymidine base (-1 Thy). The Y16F sidechain occludes the sixth octahedral coordination site.

Backbone atoms are shown as a ribbon cartoon (red) and important side chains and -1 Thy are shown as sticks (red). The PNP moiety is shown as sticks (purple) surrounded by 3 Å SA-omit density (blue cage). Residues 236-266 were disordered (small red sphere).

Figure 9. N300 inhibition.

N300 *oriT* ssDNA cleavage velocity (v_0) inhibited by PCP, CLODRO, and ETIDRO at zero, low, and high concentrations (red, orange, and yellow, respectively; concentrations indicated on the plots; error bars represent the standard errors of velocity estimation from timecourse fits; $n \geq 3$). The competitive inhibition (K_{ic}) and maximum velocity/uninhibited Michaelis constants ($V_{max}/K_m$are from nonlinear regression.

Figure 10. EC_{50} ranges.

Visual example of EC_{50} ranges calculations in **Figure 4B**. EC_{50}s are indicated for F^+ (red drop arrows) and transconjugants (yellow drop arrows) for the averaged the relative cell count curves and for ±1σ error envelope boundaries (shaded regions).

Figure 2-1 F TraI and Two-Path Model of Conjugation

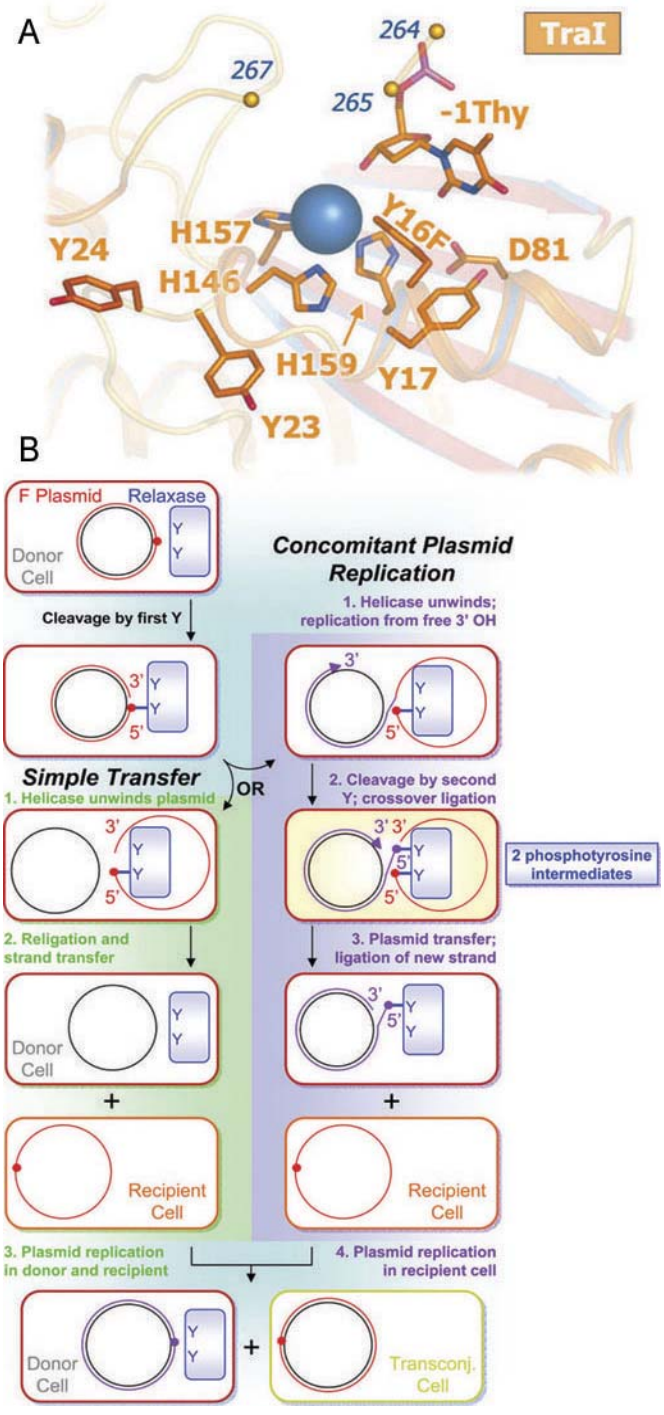


Figure 2-2 Relaxase Inhibition by PNP

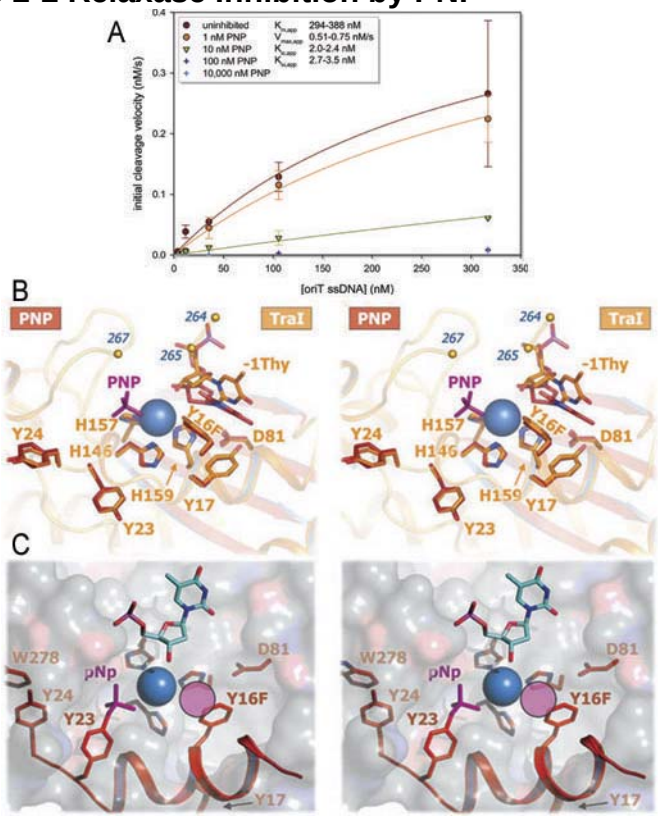


Figure 2-3 Bisphosphonates Examined for Relaxase Inhibition

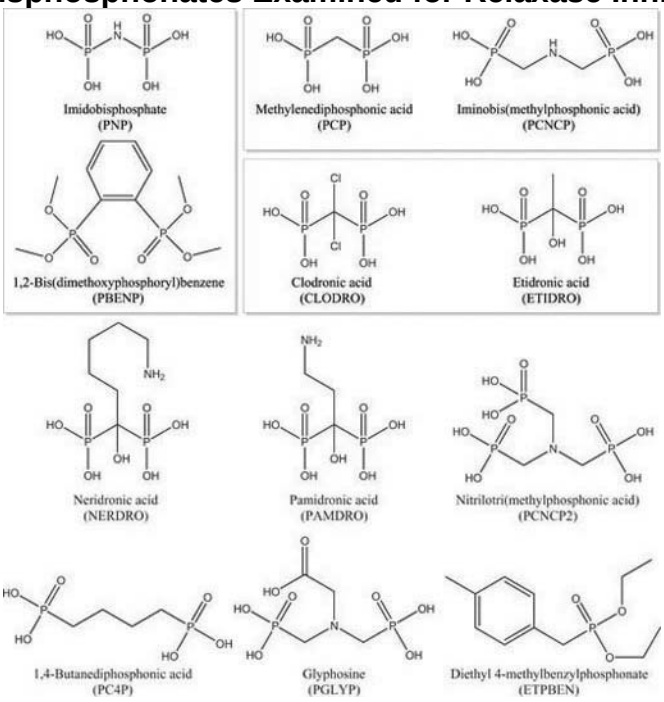


Figure 2-4 Effects of F Tral inhibitors

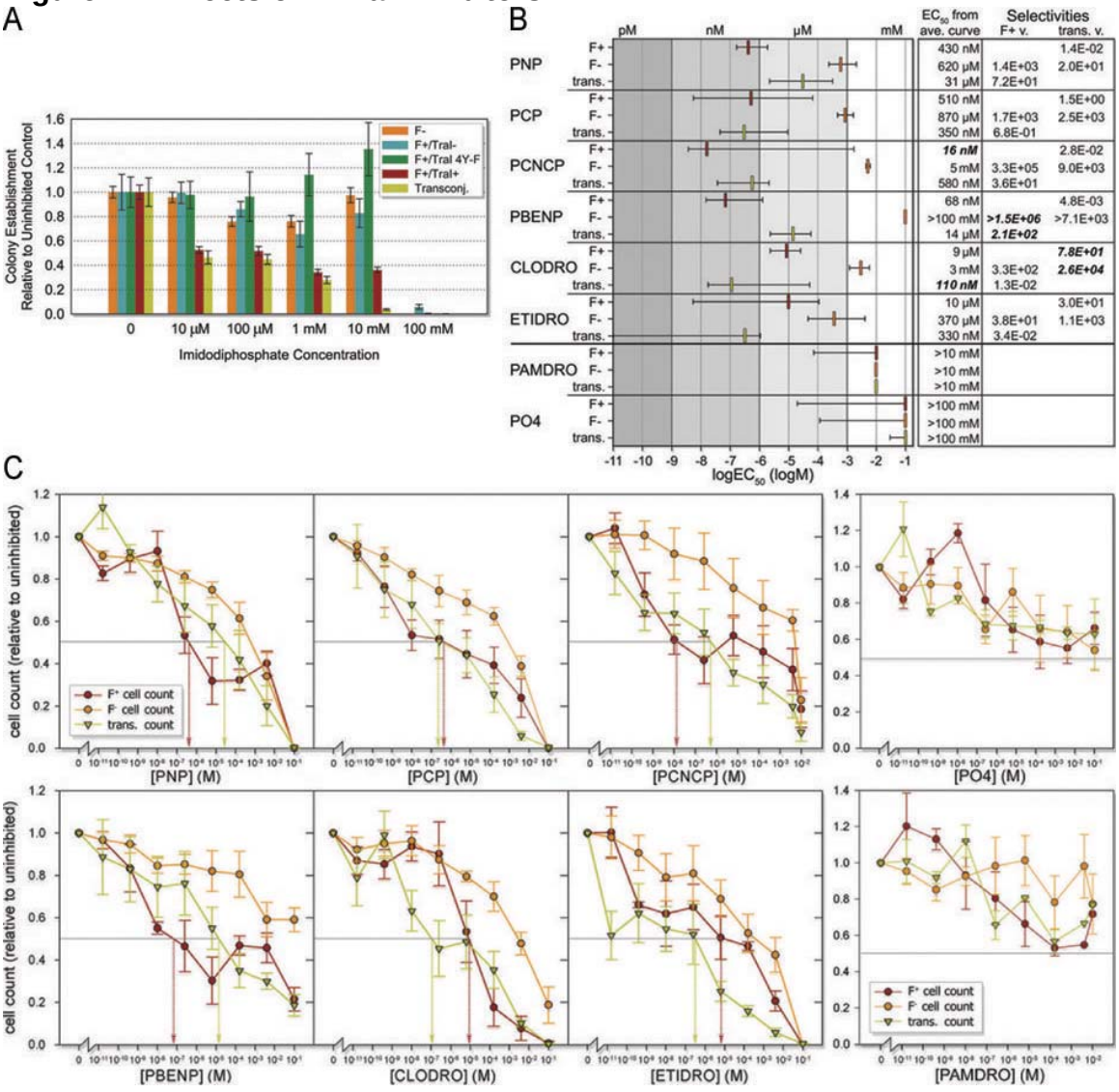


Figure 2-5 Superimposed TraI and TrwC active sites

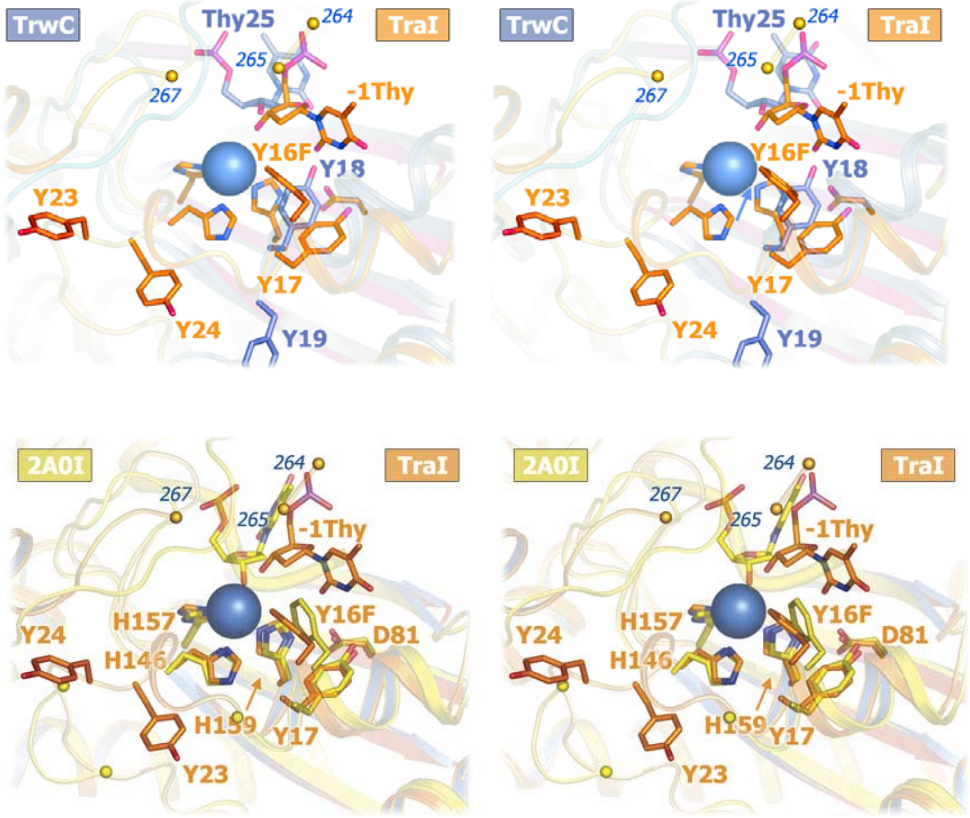


Figure 2-6 TraI 300 wild type vs. TraI 300 H159E Cleavage Activity

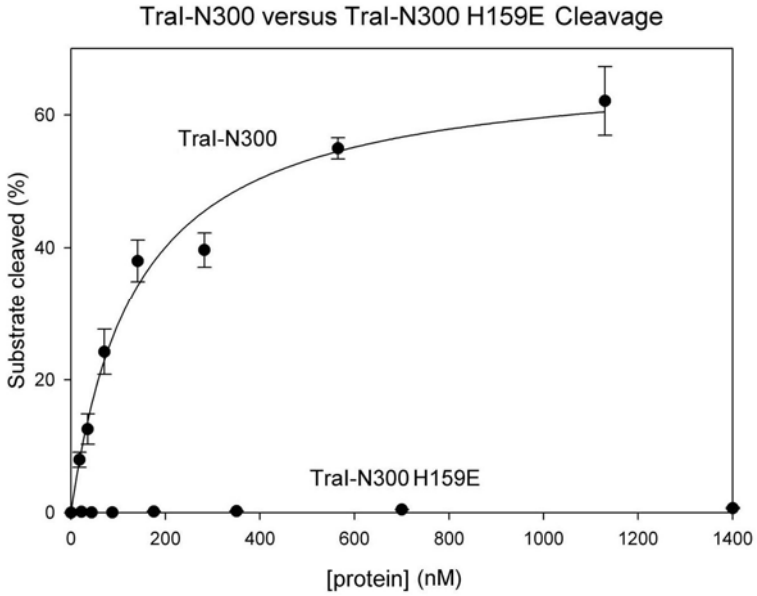


Figure 2-7 Stereo ribbon cartoon of the N300 active

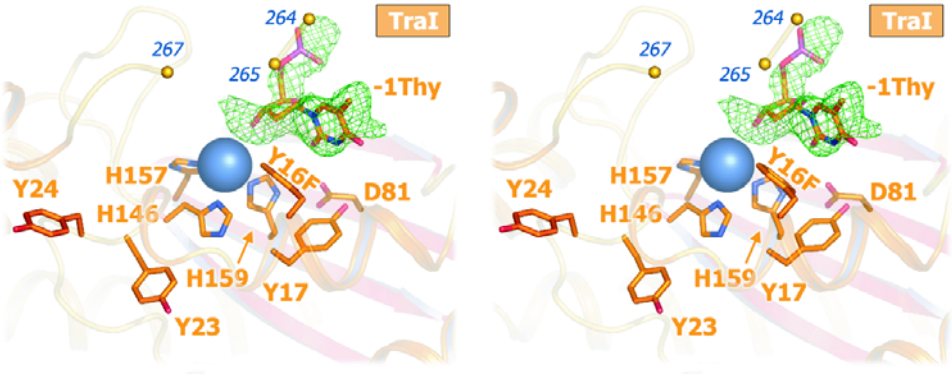


Figure 2-8 The N300 Y16F active site of the 3.0 Å structure with PNP

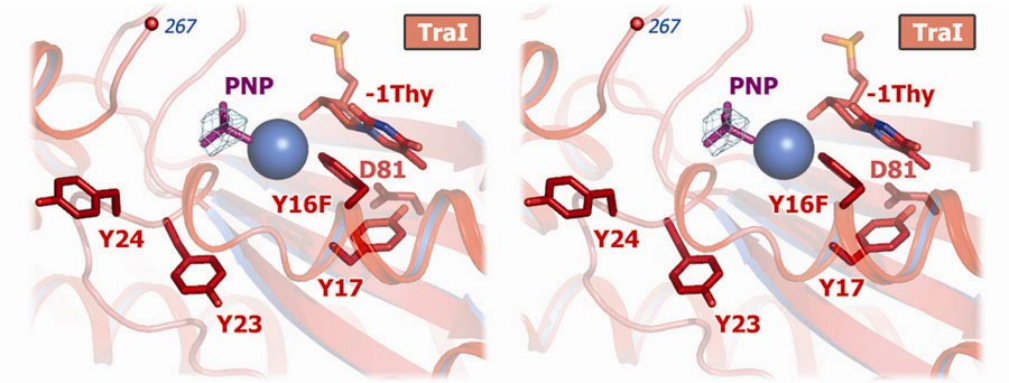


Figure 2-9 N300 inhibition

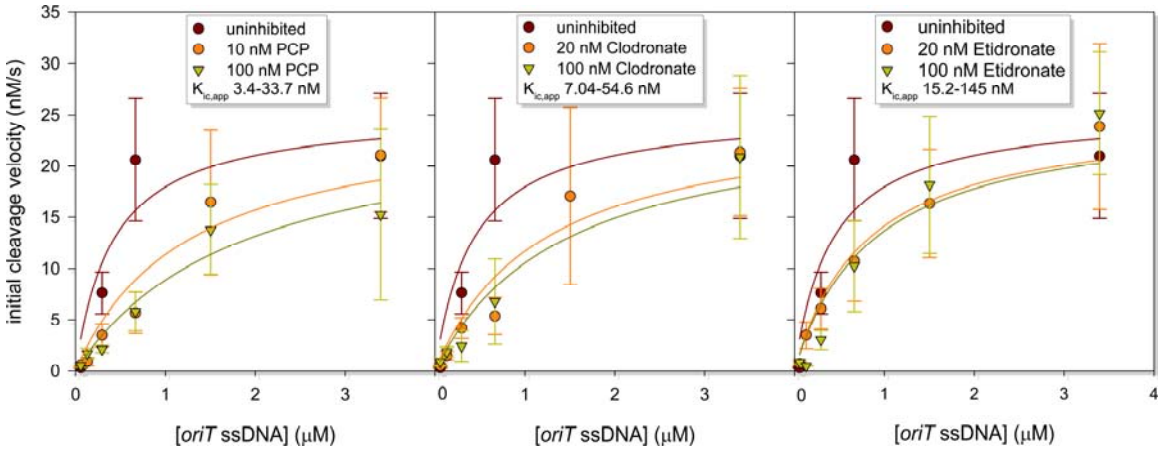


Figure 2-10 EC_{50} ranges

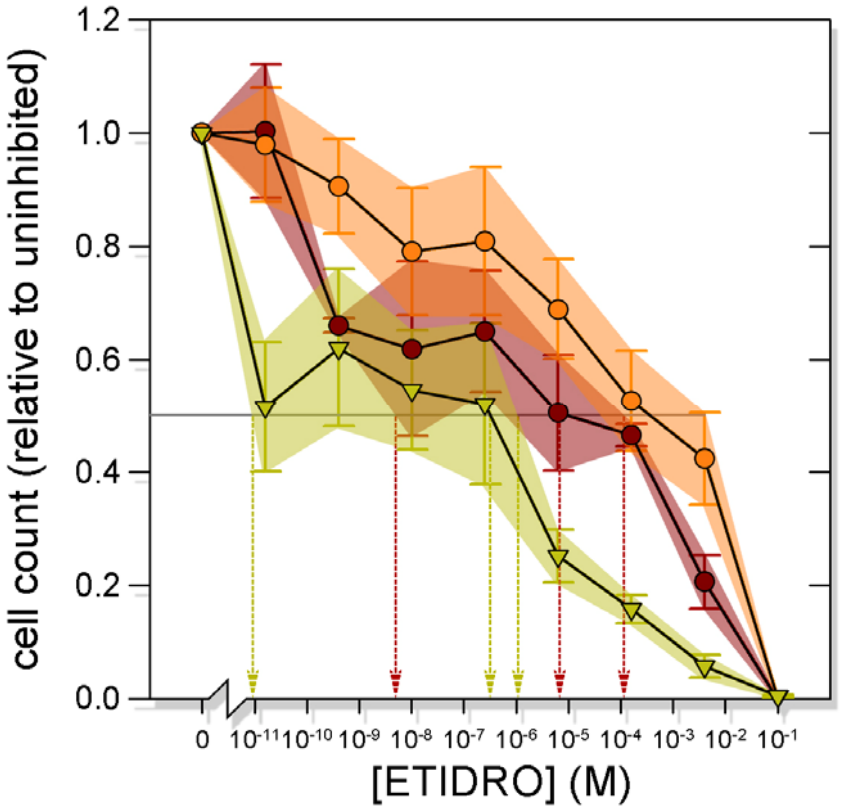


Table 2-1 Structure Statistics

Data collection and refinement	N300 - PNP	N300 + PNP
Data collection		
Space group	P $2_12_12_1$	P $2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	44.9, 88.3, 127.3	44.5, 86.3, 127.9
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	2.42-500 (2.42-2.46)	3.00-500 (3.00-3.11)
R_{sym}	9.8 (38.2)	13.5 (38.2)
I/σ	9.6	8.6
Completeness (%)	97.9 (99.3)	92.8 (72.7)
Redundancy	5.1 (5.2)	4.1 (3.9)
Refinement		
Resolution (Å)	2.42	3.00
No. reflections	19,657	10,447
R_{work} / R_{free}	21.3 / 27.0	21.4 / 31.2
No. atoms		
Protein	4142	4132
Ligand/ion	22	26
Water	157	45
B-factors		
Protein	42.4	37.2
Ligand/ion	97.7	93.5
Water	40.1	29.0
R.m.s deviations		
Bond lengths (Å)	0.0066	0.0085
Bond angles (°)	1.22	1.22

Table 2-2 Comparison of Magnesium-Binding Sites

Magnesium-Binding Proteins*		
magnesium-binding proteins	709	
backbone chelators only	51	7%
sidechain chelators	658	93%
at least one acidic chelator	518	79%
no acidic chelators	140	21%
polyphosphate binding	133	95%
other (domain/multimer interfaces)	7	5%

* Protein Data Bank crystal structures, ≤2.5 Å resolution, as of Sept. 2003.

Culled from(1); available at http://tanna.bch.ed.ac.uk/

Table 2-3 Projected Bisphosphonate Doses

Projected dose effectiveness	maximum dose (g/day) [a]	availability [b]	estimated gut concentration (mM) [c]	decrease in dilute media [d] F^+	transfer	F^-
disodium clodronate tetrahydrate	3.2	98%	5.8	93%	91%	55%
disodium etidronate	1.8	97%	4.7	95%	81%	59%

a. for a 90 kg patient, from Schering's Bonefos® and Proctor & Gamble's Didronel® product documentation.

b. 100% minus serum availability(2, 3).

c. given estimated availability and 1.5 L average gut volume.

d. 100 minute exposure, from relative cell count curves, see **Fig. 4b-4c**.

2.6 References

1. Burrus, V. & Waldor, M. K. (2004) *Res Microbiol* **155,** 376-386.

2. Lederberg, J. & Tatum, E. L. (1946) *Nature (London)* **158**.

3. de la Cruz, F. & Davies, J. (2000) *Trends Microbiol* **8,** 128-133.

4. Mazel, D. & Davies, J. (1999) *Cell Mol Life Sci* **56,** 742-754.

5. Nakamura, S., Inoue, S., Shimizu, M., Iyobe, S., & Mitsuhashi, S. (1976) *Antimicrob Agents Chemother* **10,** 779-785.

6. George, B. A. & Fagerberg, D. J. (1984) *Am J Vet Res* **45,** 2336-2341.

7. Oliva, B., Selan, L., Ravagnan, G., & Renzini, G. (1985) *Chemioterapia* **4,** 199-201.

8. Michel-Briand, Y. & Laporte, J. M. (1985) *J Gen Microbiol* **131,** 2281-2284.

9. Palomares, J. C., Prados, R., & Perea, E. J. (1987) *Chemioterapia* **6,** 256-260.

10. Molnar, J., Csiszar, K., Nishioka, I., & Shoyama, Y. (1986) *Acta Microbiol Hung* **33,** 221-231.

11. Leite, A. A., Nardi, R. M., Nicoli, J. R., Chartone-Souza, E., & Nascimento, A. M. (2005) *J Gen Appl Microbiol* **51,** 21-26.

12. Weisser, J. & Wiedemann, B. (1987) *Antimicrob Agents Chemother* **31,** 531-534.

13. Debbia, E. A., Massaro, S., Campora, U., & Schito, G. C. (1994) *New Microbiol* **17,** 65-68.

14. Spengler, G., Molnar, A., Schelz, Z., Amaral, L., Sharples, D., & Molnar, J. (2006) *Curr Drug Targets* **7,** 823-841.

15. Jung, Y. D. & Ellis, L. M. (2001) *Int J Exp Pathol* **82,** 309-316.

16. Byrd, D. R. & Matson, S. W. (1997) *Mol Microbiol* **25,** 1011-1022.

17. Pansegrau, W. & Lanka, E. (1996) *Prog Nucleic Acid Res Mol Biol* **54,** 197-251.

18. Lanka, E. & Wilkins, B. M. (1995) *Annu Rev Biochem* **64,** 141-169.

19. Llosa, M., Gomis-Ruth, F. X., Coll, M., & de la Cruz Fd, F. (2002) *Mol Microbiol* **45,** 1-8.

20. Reygers, U., Wessel, R., Muller, H., & Hoffmann-Berling, H. (1991) *Embo J* **10,** 2689-2694.

21. Matson, S. W. & Morton, B. S. (1991) *J Biol Chem* **266,** 16232-16237.

22. Traxler, B. A. & Minkley, E. G., Jr. (1988) *J Mol Biol* **204,** 205-209.

23. Lahue, E. E. & Matson, S. W. (1988) *J Biol Chem* **263,** 3208-3215.

24. Schroder, G. & Lanka, E. (2005) *Plasmid* **54,** 1-25.

25. Lawley, T. D., Klimke, W. A., Gubbins, M. J., & Frost, L. S. (2003) *FEMS Microbiol Lett* **224,** 1-15.

26. Matson, S. W., Sampson, J. K., & Byrd, D. R. (2001) *J Biol Chem* **276,** 2372-2379.

27. Byrd, D. R., Sampson, J. K., Ragonese, H. M., & Matson, S. W. (2002) *J Biol Chem* **277,** 42645-42653.

28. Traxler, B. A. & Minkley, E. G., Jr. (1987) *J Bacteriol* **169,** 3251-3259.

29. Scherzinger, E., Lurz, R., Otto, S., & Dobrinski, B. (1992) *Nucleic Acids Res* **20,** 41-48.

30. Pansegrau, W., Ziegelin, G., & Lanka, E. (1990) *J Biol Chem* **265,** 10637-10644.

31. Pansegrau, W., Schroder, W., & Lanka, E. (1994) *J Biol Chem* **269,** 2782-2789.

32. Furuya, N., Nisioka, T., & Komano, T. (1991) *J Bacteriol* **173,** 2231-2237.

33. Young, C. & Nester, E. W. (1988) *J Bacteriol* **170,** 3367-3374.

34. Srinivas, P., Kilic, A. O., & Vijayakumar, M. N. (1997) *Plasmid* **37,** 42-50.

35. Paterson, E. S., More, M. I., Pillay, G., Cellini, C., Woodgate, R., Walker, G. C., Iyer, V. N., & Winans, S. C. (1999) *J Bacteriol* **181,** 2572-2583.

36. Llosa, M., Bolland, S., & de la Cruz, F. (1994) *J Mol Biol* **235,** 448-464.

37. Greated, A., Lambertsen, L., Williams, P. A., & Thomas, C. M. (2002) *Environ Microbiol* **4,** 856-871.

38. Boer, R., Russi, S., Guasch, A., Lucas, M., Blanco, A. G., Perez-Luque, R., Coll, M., & de la Cruz, F. (2006) *J Mol Biol* **358,** 857-869.

39. Larkin, C., Datta, S., Harley, M. J., Anderson, B. J., Ebie, A., Hargreaves, V., & Schildbach, J. F. (2005) *Structure (Camb)* **13,** 1533-1544.

40. Guasch, A., Lucas, M., Moncalian, G., Cabezas, M., Perez-Luque, R., Gomis-Ruth, F. X., de la Cruz, F., & Coll, M. (2003) *Nat Struct Biol* **10,** 1002-1010.

41. Datta, S., Larkin, C., & Schildbach, J. F. (2003) *Structure (Camb)* **11,** 1369-1379.

42. Grandoso, G., Avila, P., Cayon, A., Hernando, M. A., Llosa, M., & de la Cruz, F. (2000) *J Mol Biol* **295,** 1163-1172.

43. Dressler, D. (1970) *Proc Natl Acad Sci U S A* **67,** 1934-1942.

44. Erickson, M. J. & Meyer, R. J. (1993) *Mol Microbiol* **7,** 289-298.

45. Pansegrau, W. & Lanka, E. (1996) *J Biol Chem* **271,** 13068-13076.

46. Furuya, N. & Komano, T. (2003) *J Bacteriol* **185,** 3871-3877.

47. Draper, O., Cesar, C. E., Machon, C., de la Cruz, F., & Llosa, M. (2005) *Proc Natl Acad Sci U S A* **102,** 16385-16390.

48. Llosa, M., Bolland, S., Grandoso, G., & de la Cruz, F. (1994) *J Bacteriol* **176,** 3210-3217.

49. Parker, C., Zhang, X. L., Henderson, D., Becker, E., & Meyer, R. (2002) *Plasmid* **48,** 186-192.

50. Pansegrau, W., Balzer, D., Kruft, V., Lurz, R., & Lanka, E. (1990) *Proc Natl Acad Sci U S A* **87,** 6555-6559.

51. Harding, M. M. (2004) *Acta Crystallogr D Biol Crystallogr* **60,** 849-859.

52. Matson, S. W. & Ragonese, H. (2005) *J Bacteriol* **187,** 697-706.

53. Bachmann, B. J., Low, K. B., & Taylor, A. L. (1976) *Bacteriol Rev* **40,** 116-167.

54. Smirnova, I. N., Baykov, A. A., & Avaeva, S. M. (1986) *FEBS Lett* **206,** 121-124.

55. Taylor, J. S. (1981) *J Biol Chem* **256,** 9793-9795.

56. Tomaszek, T. A., Jr. & Schuster, S. M. (1986) *J Biol Chem* **261,** 2264-2269.

57. Yount, R. G., Babcock, D., Ballantyne, W., & Ojala, D. (1971) *Biochemistry* **10,** 2484-2489.

58. Russell, R. G. (2006) *Ann N Y Acad Sci* **1068,** 367-401.

59. Brody, K. R., Hosain, P., Spencer, R. P., Hosain, F., & Wagner, H. N. (1976) *Br J Radiol* **49,** 267-269.

60. Huckell, V. F., Lyster, D. M., Morrison, R. T., & Cooper, J. A. (1985) *Clin Nucl Med* **10,** 455-462.

61. Wodnicka, M., Guarino, R. D., Hemperly, J. J., Timmins, M. R., Stitt, D., & Pitner, J. B. (2000) *J Biomol Screen* **5,** 141-152.

62. Ahmed, A. M., Kawamoto, H., Inouye, K., Hashiwata, Y., Sakaki, M., Seno, M., & Shimamoto, T. (2005) *J Med Microbiol* **54,** 867-872.

63. Tosini, F., Visca, P., Luzzi, I., Dionisi, A. M., Pezzella, C., Petrucca, A., & Carattoli, A. (1998) *Antimicrob Agents Chemother* **42,** 3053-3058.

64. Wei, Z. Q., Chen, Y. G., Yu, Y. S., Lu, W. X., & Li, L. J. (2005) *J Med Microbiol* **54,** 885-888.

65. Villikka, K., Perttunen, K., Rosnell, J., Ikavalko, H., Vaho, H., & Pylkkanen, L. (2002) *Bone* **31,** 418-421.

66. Recker, R. R. & Saville, P. D. (1973) *Toxicol Appl Pharmacol* **24,** 580-589.

67. Collaborative Computing Project, N. (1994) *Acta Crystallographica D* **50,** 760-763.

68. Jones, T., Zou, J., Cowan, S., & Kjeldgaard, M. (1991) *Acta Crystallographica A* **47,** 110-119.

69. Read, R. J. (1986) *Acta Crystallographica Section A* **42,** 140-149.

70. Brunger, A. (1998) *Acta Crystallographica D* **54,** 905-921.

71. DeLano, W. L. (2002) (DeLano Scientific, San Carlos, CA).

72. Sherman, J. A. & Matson, S. W. (1994) *J Biol Chem* **269,** 26220-26226.

73. Stern, J. C. & Schildbach, J. F. (2001) *Biochemistry* **40,** 11586-11595.

74. Burnham, K. P. & Anderson, D. R. (1998) *Model Selection and Inference* (Springer-Verlag, New York, NY).

75. Cornish-Bowden, A. & Eisenthal, R. (1974) *Biochem J* **139,** 721-730.

76. Cornish-Bowden, A. & Eisenthal, R. (1978) *Biochim Biophys Acta* **523,** 268-272.

77. Eisenthal, R. & Cornish-Bowden, A. (1974) *Biochem J* **139,** 715-720.

78. Willemoes, M., Hove-Jensen, B., & Larsen, S. (2000) *J Biol Chem* **275,** 35408-35412.

3 Further Relaxase Inhibition: Expanded Library, R100, & *in vivo* Assays

Bacterial conjugation is the primary process by which antibiotic resistance genes spread between bacterial strains, resulting in multi-drug resistance. Relaxases utilize covalent phosphotyrosine intermediates to create site- and strand-specific single stranded DNA nicks that initiate and terminate conjugative transfer. Multi-tyrosine relaxases can accommodate two simultaneous phosphotyrosine intermediates. We have recently shown that bisphosphonate compounds mimic dual-phosphotyrosine intermediates and inhibit F plasmid TraI relaxase activities *in vitro*, inhibit plasmid transfer, and selectively kill F^+ cells. Here we extend selective bisphosphonate killing to the clinical resistance-bearing F relative R100. We also expand our inhibitor library to include orders-of-magnitude more potent and selective poly-acid compounds. In gnotobiotic mice, we show that TraI inhibitors clodronate and etidronate – clinically approved bone loss treatments – cause a significant drop in F^+ load through reduction of equilibrium F^+ fractions and overall gut bacterial load. These results support the contention that

conjugative relaxase inhibition is a novel antimicrobial approach that selectively targets bacteria capable of sharing antibiotic resistance.

3.1 Introduction

The horizontal genetic transfer within and between bacterial strains are mediated by conjugative genetic elements, including those the transfer of pathogenicity factors such as antibiotic resistance (reviewed (1, 2)). Conjugative DNA transfer was first described by Lederberg and Tatum in 1946 for the *Escherichia coli* F plasmid (3). Antibiotic resistance is rapidly acquired via conjugative DNA transfer in clinical settings (reviewed (4)). Indeed, DNA conjugation generated all three documented instances of vancomycin resistant *Staphylococcus aureus* in the United States (5-9). Inhibition of conjugation of DNA conjugation or selectively killing bacteria capable of conjugation may prevent the generation and spread of multi-drug resistance in clinical settings.

We have recently described the selective inhibition of a DNA relaxase, the keystone enzyme of each conjugative system (10-12). Conjugative relaxases initiate DNA transfer with a site and strand specific nick in the transferred strand (T-strand) at the origin of transfer (*oriT*). This results in a covalent 5'-phosphotyrosine intermediate (7, 10, 13-16). The nicked T-strand is transferred from the donor cell to the recipient cell via a Type IV secretion system-mediated intercellular junction (reviewed (7, 17, 18)). Transfer is complete when the relaxase reverses the phosphotyrosine linkage, ligating and releasing the T-

strand. Many conjugative relaxases use one catalytic tyrosine (*e.g.,* relaxases from plasmid incompatibility groups IncQ, IncP, and IncI (19-22), the *Agrobacterium* inter-kindom transfer system (23), and certain conjugative transposons (24)). In the F plasmid system, however, the relaxase, is the N-terminal domain of a multifunctional protein, TraI (15, 16, 25-27). F TraI-like relaxases (*e.g.,* plasmids of incompatibility groups IncF, IncN, IncW, and *Pseudomonas* IncP9 (27-30), including many resistance-bearing plasmids) have an N-terminal constellation of two to five tyrosines nearby a chelated metal ion (31-35). Two catalytic tyrosines and a metal ion are require for maximum cleavage, ligation and transfer activity (36). Additional tyrosines may be used to resolve a replication intermediate created when replication begins at the original nick site during conjugation. This could be accomplished by creating a second nick, resulting in two simultaneous phosphotyrosine linkages. Indeed, a dual-phosphotyrosine intermediate has been trapped with suicide oligonucleotides (37). We previously predicted that this dual-phosphotyrosine intermediate would exist and used it as a starting place for inhibitor design (35).

Past studies indicated that some polycyclic chemicals, antibiotics and crude extracts inhibit of conjugation (38-44). However, these effects are mainly attributable to inhibition of bacterial growth or DNA synthesis and are independent of conjugation or conjugative transfer capability (45-48). Previously, we reported that several bisphosphonates both inhibit conjugation and selectively kill cells harboring the F plasmid (35). In this study, we expand our inhibitor

library beyond bisphosphonates to other bis-acids. We also demonstrate selective killing of cells with an F-related clinical resistance plasmid, R100. Perhaps most significantly, we show that two clinically approved bisphosphonates, etidronate and clodronate, eliminate up to 95% of F^+ *Escherichia coli* in a mammalian model at the standard human dose per unit body mass.

3.2 Results

3.2.1 In cell inhibitor library

We employ a fluorescence-based 96-well cell-enumeration assay to determine the conjugation inhibition and selective F^+ killing efficacy of a library of compounds. Seven bisphosphonate compounds were previously tested on the F plasmid system by this method (35), of which six were selective killers of F^+ *E. coli* and five of which were inhibitors of conjugation with EC50 values of <10 μM. We now expand this library with an additional 19 poly-acids, a mixture of phosphonate, carboxylate, sulfonate, and arsonate compounds. (**Figure 1)** Each was assayed for activity versus F^+ and F^- *E. coli* and against F plasmid conjugation. The dose-response curves were often complex, occasionally multiphasic, with various inflections and non-zero plateaus. Thus, half-effective concentrations (EC50s) are concentrations at which the process or population in question is reduced to 50% of uninhibited controls, with no implication of EC50 identity with any inflection point and no bearing on the relative levels of any plateaus or baselines.

At the 50% effective level, eight of these compounds were potent selective killers of F^+ cells (EC50s versus F^+ <2 μM and versus F^- >1 mM). (**Table 1**) The most potent and selective of these was SCS, with an anti-F^+ EC50 of 4.8 nM, and an F^+-versus-F^- EC50 ratio of $2.1x10^6$. Nine of the compounds were effective inhibitors of conjugation, with EC50s <6 μM. The most potent of these was SPhCOO, with an EC50 versus conjugation of 4.8 nM ($3.6x10^7$ times lower than its EC50 versus F^- cells). The method for estimating conjugation uses transconjugant counts and does not distinguish between killing of transconjugants and inhibition of the conjugative process. However, four of the compounds (citrate, SCYSCO, PCNCPM, and PCNCPE) exhibit little toxicity but permit the formation of few transconjugants, implying potent inhibition of conjugation to a degree unobserved in the previous study (35). Indeed, the highest conjugation-versus-F^+ EC50 ratio of the previous study was for clodronate at 78, while citrate has a ratio of $7.5x10^6$. Five compounds (SCS, SCCS, SBENCO, PCCCOO, and SPhCOO) were selective against both donor cells and conjugation. Three of the compounds (SNS, SCCCS, and malonate) were ineffective both at killing F^+ cells and at inhibiting conjugation. Four of the compounds (ABENA, SBENS, SCCOO, and PCCOO) were very potent killers of F^+ cells and inhibitors of conjugation, but also evidenced toxicity versus F^- cells. Thus, we have now identified four additional relaxase inhibitors with better than 20 μM EC50 versus conjugative DNA transfer, three with comparable effect on F^+ cell survival, and nine with such effects on both.

3.2.2 Lethality with a clinical resistance plasmid

Five compounds from the expanded library, SCS, CIT, PNP, PBENP, and ETIDRO, were tested against MC4100 *E. coli* cells bearing resistance plasmid R100-1 (R100-1$^+$ cells). PNP, PBENP, and ETIDRO had potent EC50 values (less than 20 μM) for killing R100-1$^+$ cells, PNP and ETIDRO less than 100 nM. SCS and CIT exhibited no anti-R100-1$^+$ effects. CIT had shown little anti-F$^+$ effect, SCS had been a potent selective killer of F$^+$ cells.

3.2.3 Decrease in total bacterial load and F$^+$ fraction

C57 BI/6J mice were administered by gavage a once daily 200 μl dose of 100 μg/kg etidronate or 180 μg/kg clodronate. These mice showed no statistically significant changes in terms of total bacterial load or of F+ fraction in fecal matter, relative to the control cohort (16 day treatment period).

The dosage was raised and the delivery method was adjusted to allow for more constant gut concentrations. Sample group mice were provided drinking water with 20 mg/ml etidronate (*n=3*) or clodronate (*n=4*). Given the approximate daily water consumption of these mice, these concentrations roughly correspond to one tenth of the maximum approved human dose per unit body mass. Within 7 days, the total fecal bacterial load did not change significantly, but for mice treated either clodronate or etidronate, the F+ fraction of remaining bacteria fell to around 60%. Over the next two days, the F+ fraction remained unchanged for

clodronate-treated mice, and rebounded to control levels for etidronate-treated mice. (**Figure 3**)

The original mice were shuffled and 10 new mice were added, doubling cohort sizes. After confirmed infection and bacterial load stabilization, sample cohorts were provided drinking water with 200 mg/ml etidronate (*n=6*) or clodronate (*n=8*), roughly equivalent to the maximum approved human dose per unit body mass. (**Figure 4**) One cage of mice (3 mice) did not respond to etidronate treatment. This cage is neglected for all future etidronate calculations (new *n=3*). Within 5 days, the total fecal bacterial load, corrected for differences in initial loads and normalized versus control, dropped to 14 ± 3% and 21 ± 4% (s.e.m.), for mice treated with etidronate or clodronate respectively. The loads varied somewhat over the remaining ten days of treatment, with mean values of 18 ± 6% and 28 ± 6% (s.d.; neglecting an anomalous spike in bacterial load on the eighth day of treatment, not shown). Likewise, the F+ fraction of remaining bacteria fell to 29 ± 7% and 21 ± 3% (s.e.m.) in the first 5 days, and stabilized at 48 ± 17% and 33 ± 9% (s.d.). This led to an initial dip in F^+ loads to 5 ± 2% (s.d.) with both drugs, and equilibrium F^+ loads of 18 ± 17% and 18 ± 14% (s.d.), for mice treated with etidronate or clodronate respectively.

3.3 Discussion

3.3.1 Qualitative SAR analysis

The complex nature of many of the *in cell* dose response curves complicates analysis of structure-activity relationships (SAR). As stated, curves occasionally multiphasic, with various inflections and non-zero plateaus. Some of these, especially for F^+ survival, may be explained as combinations of inhibitory effects and pH effects. For example, the dose response of F^+ cells against PNP has a sigmoid curve with a minimum of ~35% and inflection at ~100 nM, until the PNP concentration exceeds 4 mM and overcomes the buffering capacity of the medium. The survival curve then drops to <1% by 100 mM PNP. Multiphasic transconjugant curves generally correspond to multiphasic donor or recipient survival curves.

Our primary SAR analysis is based on effective concentrations from normalized donor, transconjugant, and recipient counts cell dose response curves (ECXs, where X = percent decrease versus control; e.g. an EC75 is the inhibitor concentration resulting in 25% of the control value). The basic TraI inhibitor framework is a poly-acid composed of four parts: two acid moieties, a bridge between the acids, and a sidechain or sidechains on the bridge. In our previous study, all poly-acid moieties were phosphonic, bridge lengths varied between 1-4 atoms, and side chains ranged from one hydrogen (PNP) to a long amidated chain (NERDRO). PBENP was unique as the one example where the sidechain cycled with the bridge and as the one example where the acidic oxygen atoms were blocked by methoxy moieties. This study expands the library

to include carboxylate, sulfonate, and arsonate poly-acids, as well as mixed-moiety poly-acids (e.g. PCCOO). Among compounds tested *in cell*, overall charges range from 0 (PBENP) to -4 (e.g. PCP).

In general, inclusion of a benzene ring, either cyclized with the bridge or on a short sidechain, increased potency versus all cell types, but most especially versus recipient cells (at the EC50 level, compare SBENS to SCCS and SPhCOO to SCCOO). The differing selectivities of ABENA and PBENP suggest that methoxylation of at least highly charged acid moieties can protect recipient cells from some lethal effects. Every potent compound with a bridge length >2 has at least one hydrogen-bonder on the bridge or sidechain (counter example: SCCCS). This may be due to increased hydrophylicity or perhaps through direct interaction of hydroxyl with the magnesium ion, analogous to the interaction of the 3' hydroxyl of the scissile thymidine, as seen in previous structures (35).

All four compounds assayed in this study that had three-atom bridges and multi-atom sidechains (PCNCPM, PCNCPE, citrate, and SCYSCO) are potent transconjugation inhibitors and ineffective at donor killing at the EC50 level. Each of these has hydrogen bond donors and acceptors on its sidechain. The three of these with hydroxylated sidechains (PCNCPM, PCNCPE, citrate, but not SCYSCO) are potent at the EC50 level. These compounds were chosen for their presumed ability to chelate the active site magnesium at all three available octahedral coordination sites. Six compounds were chosen to fulfill these

requirements in our previous study: ETIDRO, CLODRO, PAMDRO, NERDRO, PGLYP, and PCNCP2. ETIDRO and CLODRO proved to be potent both *in vitro* and *in cell*. With EC50 ratios of 30 and 78, respectively, ETIDRO and CLODRO were also the most selective among previously tested inhibitors for conjugation inhibition over donor cell killing. That said, conjugation inhibition selectivity of ETIDRO and CLODRO pale in comparison with PCNCPM, PCNCPE, or citrate (transfer versus donor killing EC50 ratios of 2500, 15,000, and 750,000, respectively). The other four previous compounds, PAMDRO, NERDRO, PGLYP, and PCNCP2, had charged and flexible sidechains (at least two free-rotating bonds), and all failed in a coarse screen as *in vitro* TraI inhibitors. We hypothesized that the charged character of the sidechains was causing some unfavorable interaction in the active site. This hypothesis was disproved by citrate. Citrate was the only compound with a charged sidechain in this study, and it was one of the most potent and selective inhibitors of tranconjugation reported herein. We thus conclude that flexible sidechains preclude TraI inhibition, at least on poly-acids with 3-atom bridges. 3-atom bridges may be the optimal poly-acid tether, allowing flexible poly-acid and hydroxyl positioning on the active site magnesium coordination sites. Compounds with longer or shorter bridges were somewhat effective at the EC50 level, but were selective for donor killing over conjugation inhibition (e.g. SGLUTS and SCOCOS).

At the EC75 level, ETIDRO, citrate, SCOCOS, ABENA, SBENS, SCCOO, and PCCOO are the only potent inhibitors thus far assayed. Only SCOCOS,

ABENA, and SCCOO have EC80 values below the 20 μM potency cutoff, all for donor cell killing, though ETIDRO, citrate, and SBENS only narrowly miss, with EC80s of 30-50 μM. ABENA qualifies as potent at the EC85 level (9.9 μM) with only SCCOO close (29 μM). Only ABENA has EC95+ values in the micromolar range, and only ABENA has an EC99 in the single-digit millimolar range (~3 mM), corresponding to less than 1 μg/ml.

The success of poly-acids with hyrdoxylated or benzylated bridges, in every metric employed, is unsurprising given their resemblance to our idealized pharmacophore. The aromatic rings were intended to interact with either the aliphatic region of the lysine 265 sidechain or with the sidechain of one of the catalytic tyrosines. The poly-acid moieties were intended to fill the octahedral coordination sites of the active site magnesium ion. The most interesting new result is the conjugative inhibition of citrate, given its physiologically innocuous nature. Future studies will focus on citrate and derivatives or analogues thereof, and on poly-acids with various aromatic rings. We are particularly interested in investigating compounds that combine aromatic rings with citrate-like hydroxylated bridges in hopes of finding compounds with the selectivity of citrate with the potency of the aromatics.

3.3.2 Bisphosphonates in other systems

Five compounds from the expanded library, SCS, CIT, PNP, PBENP, and ETIDRO, were tested against cells bearing resistance plasmid R100-1, a

derepressed derivative of R100, a close relative of the F plasmid. All of these compounds, except CIT, had exhibited potent EC50 values (less than 20 μM) for killing F^+ cells. CIT had little anti-F^+ effect, with an EC50 in the hundreds-to-thousands of micromolar range. PNP, PBENP, and ETIDRO displayed potent EC50 ranges when tested against R100-1^+ cells. The differences in EC50 values between F^+ and R100-1^+ show that there are strain and plasmid specific effects that merit further investigation.

R100 is the first clinical resistance plasmid for which selective bisphosphonate lethality has been observed. This extends the antimicrobial activity of simple bisphosphonates beyond the F plasmid model system. R100 is a close relative of the F plasmid, with approximately 98% identity between transfer regions. Studies with more distantly related conjugative systems are ongoing.

In the mouse gut, we have shown the efficacy of ETIDRO and CLODRO for reducing the F^+ bacterial fraction and for reducing the overall bacterial load when F^+ cells are present. ETIDRO and CLODRO at 200 mg/L, approximately the current approved human dose (by body mass and average daily mouse water consumption), reduced the F^+ *E. coli* load by 95±2.3% by the first timepoint. For the F plasmid system, a low dose of bisphosphonates appears to reduce the potential for the transfer of antibiotic resistance in the mammalian gut twenty-fold.

3.4 Materials and Methods

3.4.1 Bacterial strains and plasmids

F-plasmid donor cells (JS10) and recipient cells (JS4) were provided by the laboratory of Steve Matson (University of North Carolina at Chapel Hill, USA) and have been described (25). Briefly, JS10 *Escherichia coli* contain the pOX38T tetracycline resistance encoding min-F plasmid. JS4 *E. coli* have genomic streptomycin resistance. Plasmids R100 and R100-1 (in MC4100 and BL21 *E. coli* cells, respectively) were provided by the laboratory of Laura Frost (University of Alberta, Canada).

3.4.2 Fluorescent Mating & Toxicity, selection in liquid media

Inhibitor candidates were purchased from Fisher Scientific or Sigma-Aldrich (especially Sigma-Aldrich's Rare Chemical Library). Growth, mating and preparation were proceeded as previously described (35) (based originally on (25)). Briefly, saturated, antibiotic-selected overnight cultures of both donor and recipient strains were pelleted, suspended in fresh media without antibiotics, diluted 1:50 and grown at 37 ºC to an OD_{600} ~0.6. Donors, recipients and inhibitors were mixed in a 9:81:10 ratio, respectively. Final inhibitor concentrations ranged from low picomolar to high millimolar. Selection of donors, recipients and transconjugants were performed in Oxygen Biosensor™ 96-well plates (BD Biosciences). Cell concentrations were quantified on a Pherastar fluorescent plate reader (BMG) via fluorescent doubling-time analysis as previously described (49).

3.4.3 Care and handling of gnotobiotic mice

Three cohorts (I-III) of germfree inbred C57 BI/6J mice (aged 6 months) supplied by the Gnotobiotic Core at The University of North Carolina at Chapel Hill under the auspices of the Center for Gastrointestinal Biology and Disease (CGIBD) Gnotobiotic Core. Mice were inoculated with JS10 donor *E. coli* cells via direct application of pure cultures to mouth and anus and inoculation of food, or by cohabitation with mice thus infected. Mice were maintained in flexible plastic isolators at the CGIBD. Fecal bacterial populations were assayed for strain purity via MacConkey plates. Only dark colonies of lactic acid-metabolizing bacteria, such as *E. coli*, were observed on MacConkey plates. The morphology of gram stained cells from fecal samples matched that of cells from the pure JS10 inoculum. Antibiotic selection on Luria-Bertani agar (with or without 5 µg/ml of tetracycline and incubated at 37°C under aerobic conditions) was used to assay fecal samples for plasmid content. All mice were confirmed to have a numerically stable, 100% tetracycline resistant *E. coli* population within two weeks of inoculation (within nine days for all but on mouse). All experimental procedures were approved by the University of North Carolina Animal Use and Care Committee.

3.4.4 Plasmid fraction and bacterial load

After the first 10 mice were confirmed to have stable bacterial loads, sample group mice were administered a once daily 200 µl gavage of 100 µg/kg

etidronate (cohort II, *n=3*) or 180 µg/kg disodium clodronate (cohort III, *n=4*) in sterile phosphate buffered saline (PBS). Fecal samples were assayed via antibiotic selection on Luria-Bertani agar every other day for the duration of the treatment period (16 days). No statistically significant changes were observed in terms of total bacterial load or of F+ fraction, relative to the control (cohort I, *n=3*).

After a two week rest period, in which bacterial loads and F+ fractions remained constant, the delivery method and dosage were adjusted to allow for more constant gut concentrations and the mice were randomly reassigned to cohorts. Sample group mice were provided drinking water with 20 µg/ml etidronate (*n=3*) or disodium clodronate (*n=4*), approximately one tenth of the maximum approved human dose per unit body mass. Fecal samples were diluted 10^5-fold and then enumerated via the Most Probable Number dilution method (50) as estimated based on a spreadsheet-based calculator from the United States Food and Drug Administration Online Bacteriological Analytical Manual (51) .

After a one week rest period, 10 additional mice were added to the cages and the mice shuffled to create new cohorts of double the original size. After 10 days, all mice exhibited stable infection and 100% F+ fraction with no evidence of extraneous bacterial contamination. Sample group mice were then provided drinking water with 200 mg/ml etidronate (*n=6*) or disodium clodronate (*n=8*), approximately equivalent to the maximum approved human dose per unit body

mass. This treatment lasted for 15 days, during which fecal bacterial content was assayed three times. At the end of the experiment, the mice were humanely destroyed.

3.5 Figure Legends

Table 1. EC_{50}s from relative cell counts.

Colors indicate potency: orange > 20 μM; yellow = 1-20 μM; green = 10-999 nM; blue < 10 nM. Median (med.) values are calculated from averaged cell count curves, high and low values from ±1 standard deviation curves.

Figure 1. Compounds tested for this and previous studies (35).

Characters in parentheses are the abbreviations for each compound.

Figure 2. F^+ fraction during 0.1x bisphosphonate treatment.

The fraction of F^+ JS10 *E. coli* cells in the mouse gut decreases over time during treatment with either 20 mg/L ETIDRO (red bars; s.d., n = 3) or CLODRO (orange bars; s.d., n = 4), approximately one tenth the human approved dose per unit body mass, given the average volume of water consumed daily per mouse. Yellow bars represent the control group, which received vehicle only (s.d., n = 3).

Figure 3. F^+ fraction and bacterial loads during 1x bisphosphonate treatment.

A. The fraction of F^+ JS10 *E. coli* cells in the mouse gut decreases over time during treatment with either 200 mg/L ETIDRO (red bars; s.d., n = 3) or CLODRO (orange bars; s.d., n = 8), approximately the human approved dose per unit body mass. Yellow bars represent the control group, which received vehicle only (s.d., n = 6).

B. *E. coli* bacterial loads in the mouse gut during 1x ETIDRO (red bars; s.d., n = 3) or CLODRO (orange bars; s.d., n = 8), relative to control mice and corrected for differences in initial loads.

F^+ *E. coli* bacterial loads in the mouse gut during 1x ETIDRO (red bars; s.d., n = 3) or CLODRO (orange bars; s.d., n = 8), relative to control mice and corrected for differences in initial loads.

Table 3-1 EC_{50}s from relative cell count

EC50s inhbitors	med. donor	low donor	high donor	med. recip.	low recip.	high recip.	med. trans.	low trans.	high trans.
PNP	440 nM	170 nM	1.9 µM	610 µM	240 µM	2.1 mM	31 µM	2.3 µM	320 µM
PCP	600 nM	5.8 nM	71 µM	870 µM	480 µM	1.6 mM	350 nM	45 nM	9.4 µM
SNS	6.2 mM	1.1 µM	>10 mM	>10 mM	>10 mM	>10 mM	92 µM	4.9 µM	>10 mM
SCS	4.8 nM	1.1 nM	28 nM	>10 mM	1.2 µM	>10 mM	65 nM	21 nM	240 nM
SCCS	98 nM	14 nM	7.8 µM	>100 mM	4.2 mM	>100 mM	1.4 µM	49 nM	27 mM
SCOCOS	60 nM	24 nM	470 nM	6.4 mM	1.1 mM	8.4 mM	28 µM	67 nM	92 µM
PCNCP	16 nM	3.8 nM	1.7 mM	5.2 mM	4.5 mM	6.2 mM	580 nM	37 nM	2.2 µM
PBENP	73 nM	16 nM	1.4 µM	29 mM	4.6 mM	>100 mM	14 µM	2.3 µM	57 µM
SBENS	7 nM	2.2 nM	17 nM	2 µM	280 pM	16 µM	170 pM	100 pM	280 pM
ABENA	58 nM	37 nM	88 nM	160 nM	93 nM	400 nM	920 pM	390 pM	5.5 nM
SBENCO	1.2 µM	500 nM	3.4 µM	5.3 mM	1.3 mM	7.3 mM	170 nM	35 nM	1.6 µM
ETIDRO	9.8 µM	5.4 nM	110 µM	370 µM	46 µM	4.1 mM	330 nM	<10 pM	1.1 µM
CLODRO	8.6 µM	2.4 µM	26 µM	2.9 mM	1.2 mM	5.9 mM	110 nM	18 nM	53 µM
ACCOO	1.1 µM	310 pM	2 µM	>10 mM	>10 mM	>10 mM	42 µM	2.6 µM	130 µM
SCCOO	58 nM	18 nM	170 nM	3.7 µM	500 nM	31 µM	4.1 nM	1.4 nM	8.1 nM
PCCOO	230 nM	140 nM	410 nM	780 nM	180 nM	22 µM	1.1 nM	390 pM	3.2 nM
PCCCOO	380 nM	150 nM	3.5 µM	1.3 mM	79 µM	4.7 mM	4.1 nM	2.6 nM	6.8 nM
SPhCOO	200 nM	15 nM	2.2 µM	>10 mM	10 nM	>10 mM	280 pM	120 pM	840 pM
SCCCS	>100 mM	66 µM	>100 mM	>100 mM	>100 mM	>100 mM	1.2 mM	270 µM	>100 mM
SCYSCO	59 µM	3.2 µM	820 µM	12 mM	9.5 mM	14 mM	4.3 µM	70 nM	5.1 mM
SGLUTS	1.2 µM	250 nM	5.1 µM	6.4 mM	3.7 mM	9.8 mM	63 µM	140 nM	410 µM
PCNCPE	13 mM	7 mM	19 mM	13 mM	12 mM	14 mM	5.2 µM	2.4 nM	8.4 mM
PCNCPM	140 µM	32 µM	>10 mM	>10 mM	>10 mM	>10 mM	9.5 nM	1.7 nM	620 nM
CIT	290 µM	14 µM	610 µM	1.4 mM	870 µM	2.7 mM	390 pM	89 pM	8 nM
PAMDRO	>10 mM	72 µM	>10 mM	>10 mM	>10 mM	>10 mM	>10 mM	>10 mM	>10 mM
MAL	83 µM	19 µM	7.1 mM	6 mM	4.2 mM	8 mM	66 µM	140 nM	8 mM
PO4	>100 mM	20 µM	>100 mM	>100 mM	120 µM	>100 mM	>100 mM	30 mM	>100 mM

Figure 3-1 Compounds tested for this and previous studies

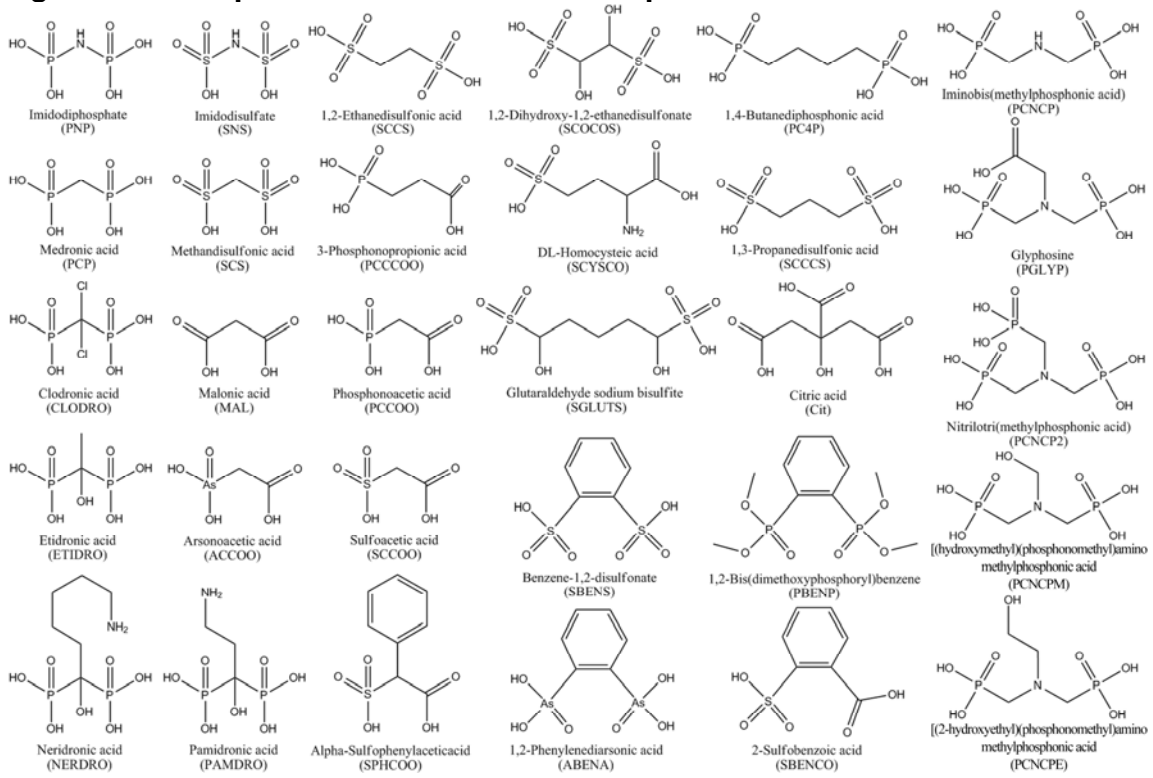


Figure 3-2 F^+ fraction during 0.1x bisphosphonate treatment

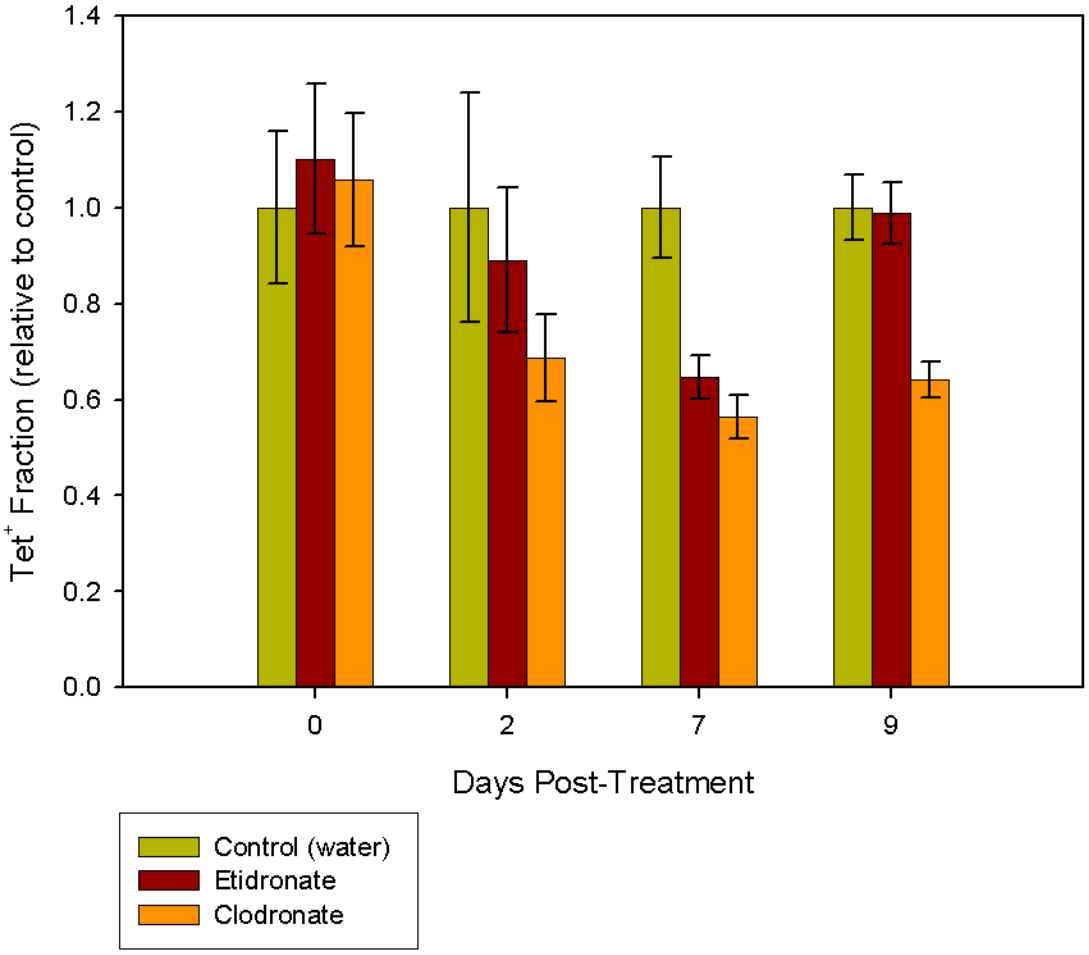
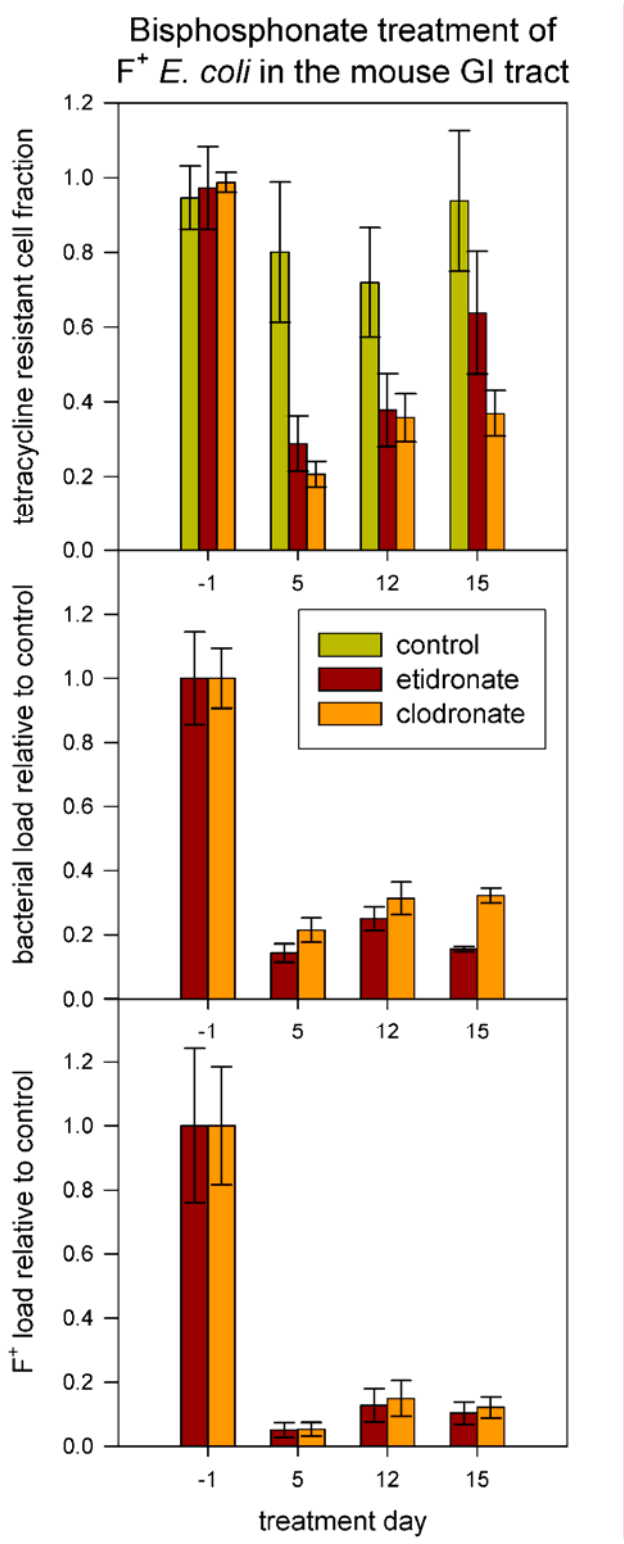


Figure 3-3 F^+ fraction and bacterial loads during 1x bisphosphonate treatment



3.6 References

1. de la Cruz, F. & Davies, J. (2000) *Trends Microbiol* **8,** 128-133.

2. Burrus, V. & Waldor, M. K. (2004) *Res Microbiol* **155,** 376-386.

3. Lederberg, J. & Tatum, E. L. (1946) *Nature (London)* **158**.

4. Mazel, D. & Davies, J. (1999) *Cell Mol Life Sci* **56,** 742-754.

5. Weigel, L. M., Donlan, R. M., Shin, D. H., Jensen, B., Clark, N. C., McDougal, L. K., Zhu, W., Musser, K. A., Thompson, J., Kohlerschmidt, D*., et al.* (2007) *Antimicrobial agents and chemotherapy* **51,** 231-238.

6. Weigel, L. M., Clewell, D. B., Gill, S. R., Clark, N. C., McDougal, L. K., Flannagan, S. E., Kolonay, J. F., Shetty, J., Killgore, G. E., & Tenover, F. C. (2003) *Science (New York, N.Y* **302,** 1569-1571.

7. Llosa, M., Gomis-Ruth, F. X., Coll, M., & de la Cruz Fd, F. (2002) *Molecular microbiology* **45,** 1-8.

8. (2002) *Mmwr* **51,** 565-567.

9. (2004) *Mmwr* **53,** 322-323.

10. Byrd, D. R. & Matson, S. W. (1997) *Molecular microbiology* **25,** 1011-1022.

11. Pansegrau, W. & Lanka, E. (1996) *Prog Nucleic Acid Res Mol Biol* **54,** 197-251.

12. Lanka, E. & Wilkins, B. M. (1995) *Annu Rev Biochem* **64,** 141-169.

13. Reygers, U., Wessel, R., Muller, H., & Hoffmann-Berling, H. (1991) *The EMBO journal* **10,** 2689-2694.

14. Matson, S. W. & Morton, B. S. (1991) *The Journal of biological chemistry* **266,** 16232-16237.

15. Traxler, B. A. & Minkley, E. G., Jr. (1988) *Journal of molecular biology* **204,** 205-209.

16. Lahue, E. E. & Matson, S. W. (1988) *The Journal of biological chemistry* **263,** 3208-3215.

17. Schroder, G. & Lanka, E. (2005) *Plasmid* **54,** 1-25.

18. Lawley, T. D., Klimke, W. A., Gubbins, M. J., & Frost, L. S. (2003) *FEMS Microbiol Lett* **224,** 1-15.

19. Scherzinger, E., Lurz, R., Otto, S., & Dobrinski, B. (1992) *Nucleic Acids Res* **20,** 41-48.

20. Pansegrau, W., Ziegelin, G., & Lanka, E. (1990) *J Biol Chem* **265,** 10637-10644.

21. Pansegrau, W., Schroder, W., & Lanka, E. (1994) *J Biol Chem* **269,** 2782-2789.

22. Furuya, N., Nisioka, T., & Komano, T. (1991) *J Bacteriol* **173,** 2231-2237.

23. Young, C. & Nester, E. W. (1988) *J Bacteriol* **170,** 3367-3374.

24. Srinivas, P., Kilic, A. O., & Vijayakumar, M. N. (1997) *Plasmid* **37,** 42-50.

25. Matson, S. W., Sampson, J. K., & Byrd, D. R. (2001) *The Journal of biological chemistry* **276,** 2372-2379.

26. Byrd, D. R., Sampson, J. K., Ragonese, H. M., & Matson, S. W. (2002) *The Journal of biological chemistry* **277,** 42645-42653.

27. Traxler, B. A. & Minkley, E. G., Jr. (1987) *J Bacteriol* **169,** 3251-3259.

28. Paterson, E. S., More, M. I., Pillay, G., Cellini, C., Woodgate, R., Walker, G. C., Iyer, V. N., & Winans, S. C. (1999) *J Bacteriol* **181,** 2572-2583.

29. Llosa, M., Bolland, S., & de la Cruz, F. (1994) *J Mol Biol* **235,** 448-464.

30. Greated, A., Lambertsen, L., Williams, P. A., & Thomas, C. M. (2002) *Environ Microbiol* **4,** 856-871.

31. Boer, R., Russi, S., Guasch, A., Lucas, M., Blanco, A. G., Perez-Luque, R., Coll, M., & de la Cruz, F. (2006) *J Mol Biol* **358,** 857-869.

32. Larkin, C., Datta, S., Harley, M. J., Anderson, B. J., Ebie, A., Hargreaves, V., & Schildbach, J. F. (2005) *Structure (Camb)* **13,** 1533-1544.

33. Guasch, A., Lucas, M., Moncalian, G., Cabezas, M., Perez-Luque, R., Gomis-Ruth, F. X., de la Cruz, F., & Coll, M. (2003) *Nat Struct Biol* **10,** 1002-1010.

34. Datta, S., Larkin, C., & Schildbach, J. F. (2003) *Structure (Camb)* **11,** 1369-1379.

35. Lujan, S. A., Guogas, L. M., Ragonese, H., Matson, S. W., & Redinbo, M. R. (2007) *Proceedings of the National Academy of Sciences of the United States of America* **104,** 12282-12287.

36. Grandoso, G., Avila, P., Cayon, A., Hernando, M. A., Llosa, M., & de la Cruz, F. (2000) *Journal of molecular biology* **295,** 1163-1172.

37. Gonzalez-Perez, B., Lucas, M., Cooke, L. A., Vyle, J. S., de la Cruz, F., & Moncalian, G. (2007) *The EMBO journal* **26,** 3847-3857.

38. Nakamura, S., Inoue, S., Shimizu, M., Iyobe, S., & Mitsuhashi, S. (1976) *Antimicrob Agents Chemother* **10,** 779-785.

39. George, B. A. & Fagerberg, D. J. (1984) *Am J Vet Res* **45,** 2336-2341.

40. Oliva, B., Selan, L., Ravagnan, G., & Renzini, G. (1985) *Chemioterapia* **4,** 199-201.

41. Michel-Briand, Y. & Laporte, J. M. (1985) *J Gen Microbiol* **131,** 2281-2284.

42. Palomares, J. C., Prados, R., & Perea, E. J. (1987) *Chemioterapia* **6,** 256-260.

43. Molnar, J., Csiszar, K., Nishioka, I., & Shoyama, Y. (1986) *Acta Microbiol Hung* **33,** 221-231.

44. Leite, A. A., Nardi, R. M., Nicoli, J. R., Chartone-Souza, E., & Nascimento, A. M. (2005) *J Gen Appl Microbiol* **51,** 21-26.

45. Weisser, J. & Wiedemann, B. (1987) *Antimicrob Agents Chemother* **31,** 531-534.

46. Debbia, E. A., Massaro, S., Campora, U., & Schito, G. C. (1994) *New Microbiol* **17,** 65-68.

47. Spengler, G., Molnar, A., Schelz, Z., Amaral, L., Sharples, D., & Molnar, J. (2006) *Curr Drug Targets* **7,** 823-841.

48. Jung, Y. D. & Ellis, L. M. (2001) *Int J Exp Pathol* **82,** 309-316.

49. Wodnicka, M., Guarino, R. D., Hemperly, J. J., Timmins, M. R., Stitt, D., & Pitner, J. B. (2000) *J Biomol Screen* **5,** 141-152.

50. Oblinger, J. L. & Koburger, J. A. (1975) *J. Milk Food Technol.* **38,** 540-545.

51. Garthright, W. E. & Blodgett, R. J. (2003) *Food Microbiology* **20,** 439-445.

4 Advent of TraI: Phylogenetics of Relaxase Evolution

Antibiotic resistance genes spread between bacterial strains primarily through bacterial conjugation. This can create multi-drug resistant bacterial strains. Relaxases create site- and strand-specific single stranded DNA nicks, via covalent phosphotyrosine intermediates, which initiate and terminate conjugative transfer. We use phylogenetic analyses of bacterial relaxases to show the origins of these clinically important enzymes and to support our previously proposal that multi-tyrosine relaxases resolve the products of concomitant DNA replication.

4.1 Introduction

Antibiotic resistance genes and other virulence factors are spread within and between bacterial strains through horizontal gene transfer (reviewed (1)). Horizontal transfer in bacteria is accomplished primarily through conjugation (reviewed (2)), as first described for the *Escherichia coli* F plasmid by Lederberg and Tatum in 1946 (3). In clinical settings, antibiotic resistance is acquired with a rapidity that is critically dependent on conjugative DNA transfer (reviewed (4)), as was the case for all documented instances of vancomycin resistant

Staphylococcus aureus in the United States (5-9). Either the inhibition of conjugation or the selective killing of conjugation capable bacteria in clinical settings would be a highly desirable suppressive of the generation and spread of multi-drug resistance.

The details of the diverse Mob family relaxase mechanisms are poorly understood. Some relaxases possess one catalytic tyrosine residue (TraA and VirD2 of *Rhizobia* and *Agrobacteria* pTi plasmids; TraI of plasmid RP4 IncQ) while others have a constellation of conserved tyrosines (TraA of *Corynebacteria* plasmids; TrwC of plasmid R388 IncW; TraI of plasmid R100 IncF plasmids). The IncF/W constellation of two tyrosine pairs follows a pattern of Yy-X_{3-26}-YY, where 'y' denotes either a tyrosine or a phenylalanine in some IncW relaxases (10). The tyrosine pairs may be responsible for multiple transesterifications similar to the rolling-circle model described for bacteriophage phi X174 (11-13). During the nicking (transesterification) reaction, a catalytic tyrosine forms a covalent bond with a phosphate of the nicked strand (14, 15). Mutant data with R388 TrwC show that all four tyrosines show varying degrees of capability for creating the initial nick (16). We have suggested that the second tyrosine set is responsible - via a dual phosphotyrosine intermediate - for resolution of a new nick site created by concurrent replication (17). Though this intermediate was recently observed (18), we undertook a phylogenetic analysis of bacterial relaxase sequences both to find the origins of multi-tyrosine relaxases and in the hope that the order of evolutionary innovations would either support or refute our hypothesis that

multiple tyrosines are needed for the resolution of concurrently replicated nick sites.

4.2 Results

4.2.1 Alignment of diverse bacterial relaxases

Homology searches, based on seed sequences for one single-tyrosine relaxase (plasmid R64 NikB) and one multi-tyrosine relaxase (*Brevibacterium linens* BL2 TraA), returned, after manual curation, over 130 relaxases sequences. These included sequences from every major bacterial plasmid incompatibility class, and chromosome and transposons sequences. Sequences were recovered from soil-, deep sea-, animal- and plant-associated gram positive and gram negative strains, including plant, wildlife, livestock, and human pathogens.

4.2.2 Relaxase phylogenetics

Assuming that the outgroup for each tree (**Figure 1**) was the TraI relaxase of IncP plasmid R751, single-tyrosine relaxases appear to be basal. Multi-tyrosine active sites appear to be monophyletic, with the *Streptomyces* relaxases pFP11.6c and pFP1.1 as the basal taxa. The multi-tyrosine clade is a sister group to the a cluster of IncQ-like relaxases, including the IncQ-like TraA relaxases of *Agrobacterium* virulence plasmids and relaxases from true IncQ plasmids, such as MobA of plasmid RSF1010. This makes the IncQ-like relaxases polyphyletic, with relaxases from IncQ-like plasmids pTC-F14, pTC-

FC2, pRAS3.1 and pRAS3.2 more closely related to IncP-6/G/U, IncL/M, and IncI1 relaxases than to true IncQ relaxases. Relaxases with carboxy-terminal helicase domains make up a paraphyletic group composed of the *Legionella/Rhizobial* TraA clade and the multi-tyrosine clade, but excluding the true IncQ relaxases and their closest relatives.

Treating Gram stain ability of the host cell as a character state, the most parsimonious tree, rooted with R751 TraI, has seven state changes and suggests a gram positive host for the last common ancestor of IncP-like and multi-tyrosine relaxases. The closest homologue to the relaxase-fused helicase domains is RecD, part of the RecBCD complex of Gram negative bacteria. Thus the last common ancestor of all helicase-fused relaxases must have come from a Gram negative cell. The most parsimonius tree that can incorporate this knowledge has eight state changes and also implies a Gram positive last common ancestor. The chosen root is from a gram positive source, as are the basal groups for both the muti-tyrosine clade and the helicase/IncQ clade.

4.2.3 Motif I duplication

Streptomyces relaxases pFP11.6c and pFP1.1 bear direct repeats near their amino-termini, as detected by European Molecular Biology Laboratory - European Bioinformatics Institute (EMBL-EBI) online Rapid Automatic Detection and Alignment of Repeats (RADAR). These repeats duplicate relaxase Motif I, including the catalytic tyrosine.

4.3 Discussion

4.3.1 Ancestral relaxase host

If Gram stains are taken as a two-character state, parsimony suggests that the common ancestor of Clade 0 (all relaxases studied herein) resided in Gram positive cells. The parsimony evidence for this is relatively weak, as several trees incorporating a Gram negative root fall short of the minimum step count by only one or two steps. Given the broad range of many plasmids, it may be just as likely that the last common ancestor resided in both Gram positive and Gram negative strains.

Relaxases Clades 8 and 6 have fused carboxyl-terminal helicase domains. The closest non-relaxase relative of these helicases is RecD, of the Gram negative DNA repair complex RecBCD. Thus the common ancestor of Clades 8 and 6 must have resided in a gram negative bacterium. The distribution of taxa in Clade 2 also suggests that the ancestors at nodes 2, 4, 5, 6, 10, and 11 inhabited hosts that lived either in soil, on plants or plant detritus, or both.

Helicase fusion occurred at Node 4, implying helicase deletion at Node 7. This was unexpected and may be an artifact of maximum parsimony calculations using wildcard alignment gaps. Far more likely is that Clades 6 and 8 are sister groups that originally diverged from the IncQ MobA relaxases of Clade 7 after helicase fusion.

4.3.2 Evolution of multi-tyrosine relaxases

Motif I bifurcation occurred through partial gene duplication after helicase fusion, apparently in a *Streptomycete* host or some other Gram positive soil dweller. Duplication resulted in a chain of more than thirty residues between the tyrosine pairs, corresponding to a loop region in structures of F TraI and R388 TrwC (Clade 9) (17-21). Remnants of the long loop may be found in extant relaxases on plasmids pFP11.6c and pFP1.1, still with signs of the original duplication. This chain was paired down in subsequent relaxases: to 18-22 residues in Clade 10; to 15-25 in pNAC3 TraA; and then to 3-5 in the rest of Clade 9 (except for one clade of mostly deep sea *Vibrio* relatives (22)). The rapid reduction of this loop may suggest an advantage to stiffening the loop or clearing residues from the active site. Stiffening or simply shortening of the loop may aid in a concerted conformational change encompassing all of the tyrosines, as previously suggested for F TraI (17). It would be interesting to see the effects of insertion of a large loop in this region of F TraI or R388 TrwC, or conversely a loop deletion in one of the pFP relaxases.

4.3.3 Multiple tyrosines and fused helicases

Fixation of N-terminal secondary tyrosines occurred only after helicase fusion, which may imply an advantageous connection between helicase fusion and multiple tyrosines. Secondary tyrosines proximal to N-termini are common in Clade 7 IncQ-like relaxases and Clade 8 helicase-fused relaxases, but these

tyrosines are interchangeable with other residues, primarily hydrophobic residues. The advantage conferred by the bifurcated tyrosine constellation may be related to the speed of DNA unwinding and/or strand transfer. Conjugative helicases are among the fastest and most processive monomeric helicases known (approximately 1100 bp/s, from F TraI (23)). Assuming that the motive force for unwinding in non-helicase-fused systems is provided by the DNA replication machinery (approximately 300 bp/s, from phage T7 replication (24)), plasmid transfer for helicase-fused systems may proceed 3-4 times faster. For conjugation, faster unwinding results in faster transfer, which in turn results in fewer partial transfers due to mating pair shearing. Selective advantage results because more complete transfers mean greater proliferation.

Faster unwinding also has disadvantages. If replication begins at the nick site during transfer, then a second nick site is formed and must be resolved by a second nicking event before transfer termination. One obvious solution is to use a second relaxase to resolve the new nick site, and this may be the solution used in single-tyrosine systems. However, rapid unwinding and transfer (e.g. ~ 90 seconds for the F plasmid) allows little time for resolution of the new nick site. We have suggested that the second tyrosine set is responsible for resolution of a new nick site via a dual phosphotyrosine intermediate (17). This intermediate was recently trapped and observed using suicide oligonucleotides (18). The same study showed that the second nick was dependent upon the formation of a specific secondary DNA structure just upstream of the nick site, which the

relaxase can recognize and bind (19). Since the relaxase domain is maintained proximal to the unwinding fork by the fused helicase, there is the potential for the relaxase to rapidly recognize and process the new nick site as it passes through the helicase domain. In this sense, helicase fusion was a permissive event that bestowed a selective advantage on fixation of secondary N-terminal tyrosines.

Whatever the selective advantage gained by the combination of multiple tyrosines with helicase fusion, it has allowed the multi-tyrosine relaxases to expand beyond the soil dwelling hosts of pFP11.6c and pFP1.1 and the plant pathogen niche inhabited by Clade 8 (these helicase-fused *tra* relaxases are used for interbacteria transfer, and should not be confused with *vir* relaxases from the same strains, which are responsible for transfer between bacteria and plants and are not discussed here (25)). This move from soil and plant dwellers resulted in many of resistance-bearing conjugative plasmids that plague modern medicine.

4.4 Materials and Methods

4.4.1 Initial homology searches

Initial homology searches were made with Psi-BLAST (26). Single-tyrosine relaxase sequences were acquired with the amino acid sequence of plasmid R64 NikB (NCBI accession BAA78021) amino-terminal region, through relaxase Motif III (residues 1-167) as the search target and default Psi-BLAST settings. 500 hits were accepted. Sequences with expectation values (E-values) higher than 1e-10

were culled manually and sequences lacking canonical relaxase Motifs I and III were rejected. All sequences with canonical relaxase Motifs I and III were accepted regardless of E-value. Psi-BLAST iterations 1-4 yielded sequences with annotated *NikB*, *VirD2*, Inc P type *TraI*, *MobA*, and *Rlx*. Round five yields the first sequence annotated *TraA*. The hit list was expanded to 1000 sequences. Iterations 6 and 7 added sequences annotated as *pre*, and *recombinase*. The hit list was expanded to allow 1500 sequences, but no additional sequences met E-value or Motif criteria. The search was terminated and all sequences with E-values of 7e-10 or less were collected. (**Table 1**)

Multiple-tyrosine relaxase sequences were acquired with the peptide sequence of the *Brevibacterium linens* BL2 TraA (NCBI accession ZP_00379389) amino-terminal region, through relaxase Motif III (residues 1-174) as the search target and default Psi-BLAST settings. 500 hits were accepted. Sequences with expectation values (E-values) higher than 1e-10 were culled manually and sequences lacking canonical relaxase Motifs I and III were rejected. All sequences with canonical relaxase Motifs I and III were accepted regardless of E-value. Psi-BLAST iterations 1-4 yielded sequences with such annotations as *Corynebacterium* type *TraA*, *Agrobacterium* type *TraA* Agro, and IncF type *TraI*. Round five returned additional sequences that met E-value and Motif criteria. The search was terminated and all sequences with E-values of 4e-10 or less were collected from the National Center for Biotechnology Information (NCBI) protein sequence database.

4.4.2 Semi-automated sequence annotation

Full NCBI protein reports, in concatenated form, were parsed with a short custom Perl script. The parsed headers and protein sequences were entered into a custom Microsoft Excel 2003 spreadsheet that further parsed the Perl output and generated hyperlinks to each report page. The hyperlinks were used to access reports with ambiguous or incomplete annotation to allow manual annotation. Annotation uses the following format:

a_b_c_d y

where

a = gene (from NCBI report entry */gene* or manual)

b = locale (NCBI report entry */plasmid*, */locus_tag*, or manual)

c = organism (NCBI report entry */organism* or manual)

d = strain note (manual notation)

y = accession #

Locale (b) is noted with the following formats:

Ch_# = supposed chromosomal locus #

Ch_ = supposed chromosomal locus unknown

Chr# = locale chromosome #

Where known, the organism (c) is denoted with a five letter concatenation of the first three letters of the genus name and the first two of the species (e.g. *Escherichia coli* is denoted as Escco).

4.4.3 Initial multiple sequence alignment

Alignments were constructed in ClustalX version 1.80 (27). The initial alignment (250 sequences) was made as per BLAST (26) defaults (Blossum matrix series, gap opening/extension penalties set to 11/1). Alignment distances were used to construct a Neighbor-Joining tree from the distances in the .dnd file generated with the alignment. The tree subjected to 10 rounds of bootstrapping with ClustalX (cursory: gaps allowed and no multi-substitution correction). The sequences were divided into clusters (**Table 2**) by bootstrap value, where a cluster was defined as a maximal set of sequences with bootstrap values greater then 5 (out of 10) for all intervening nodes.

4.4.4 Tree building

Alignments created in ClustalX were edited and columns with gaps stripped in BioEdit (28). Identical sequences were combined in DAMBE. Trees were constructed in ClustalX, Protpars (Bioedit), and DAMBE (29). Trees were bootstrapped, where applicable, in ClustalX, Phylip, and DAMBE. Branches with bootstrap values of less than 50 (out of 100) were culled. Trees were visualized in Treeview (30) and colorized in Adobe Photoshop.

4.5 Special Acknowledgement

I would like to thank Daniel Timothy Winston for writing the PubMed gene report initial parser script.

4.6 Figure Legends

Figure 1. Maximum parsimony phylogram of bacterial relaxases.

This phylogram was generated from an alignment of the full amino acid sequences of 251 relaxases. Each n-residue gap in the alignment was assigned as one gap character and n-1 wildcard characters. All alignments used secondary structure masking based on F TraI and R388 TrwC structures. Important nodes are numbered for callouts in the text. Branch and title colors indicate Gram satin character, red meaning Gram positive, yellow for Gram negative, and cyan indicating cyanobacteria. Parsimony was used to approximate the Gram character for ancestral nodes. All helicase-fused bacterial relaxases branch from nodes 6 and 8 and all multi-tyrosine relaxases from node 6.

Table 1. ClustalX alignment input variables.

Table 2. Accession numbers and titles of relaxase sequences.

Figure 4-1 Maximum parsimony phylogram of bacterial relaxases

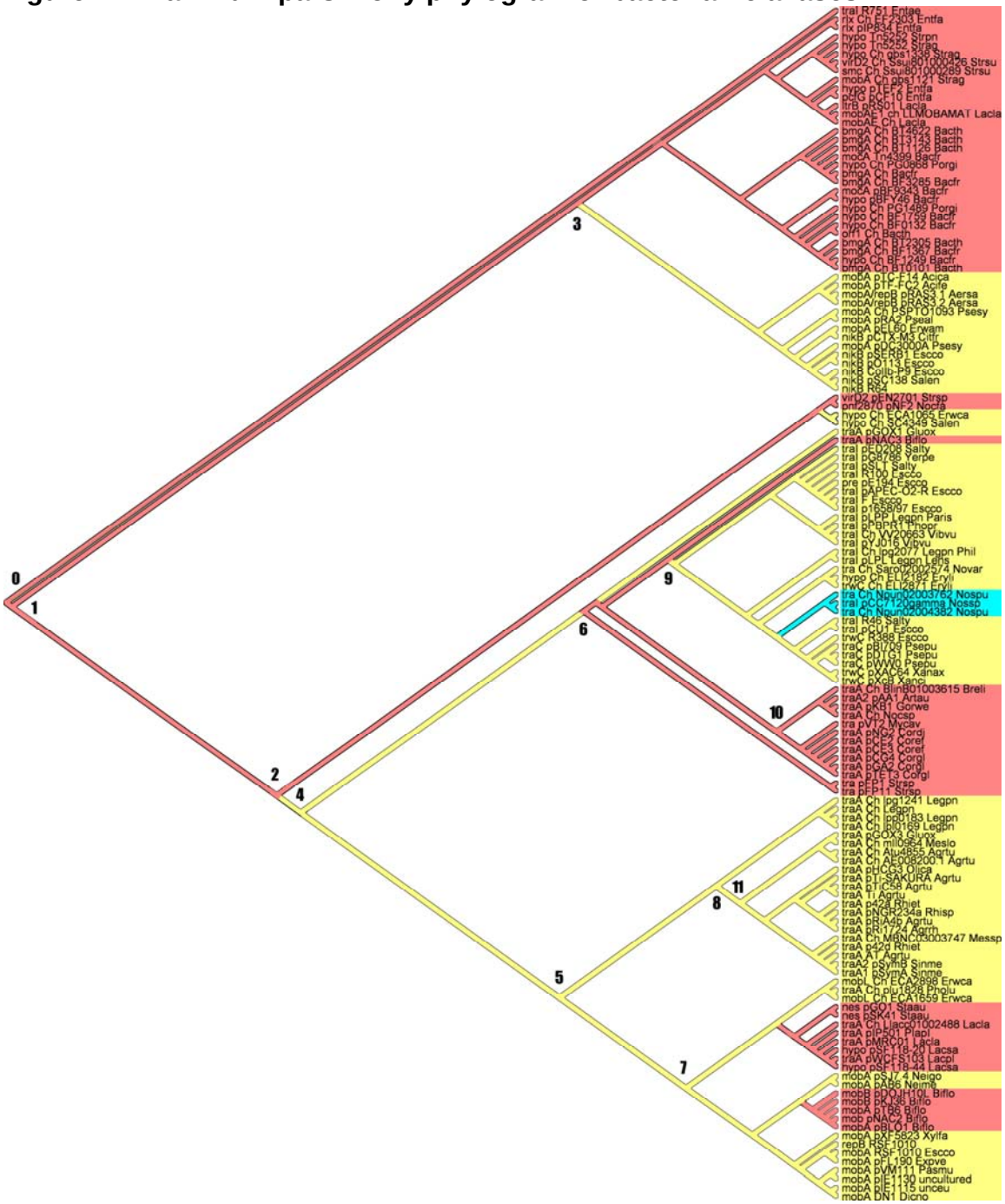


Table 4-1 ClustalX alignment input variables

Preliminary align. #	from cluster	alignment mode	matrix series	gap penalties opening	gap penalties extension	2° structure masking	manual correction?	Note
1	IncF	profile	BLOSUM	4	1	1P4D	n	
2	IncWP9	profile	BLOSUM	4	1	1QX0	n	
3	IncX	multiple	BLOSUM	6	1		y	
4	IncXb	multiple	BLOSUM	6	1		y	
5	IncQ	multiple	Gonnet	10	0.2		y	
6	IncQb	multiple	Gonnet	10	0.2		y	
7	IncPII	multiple	Gonnet	10	0.2		n	
8	GposQ	multiple	Gonnet	10	0.2		y	
9	Pre	multiple	Gonnet	10	0.2		y	
10	Cory	multiple	Gonnet	10	0.2		n	
11	IncI	multiple	Gonnet	10	0.2		y	
12	TiTra	multiple	Gonnet	10	0.2		y	
13	Gor-Str	multiple	Gonnet	10	0.2		y	shuffled tra_pFP1_Strsp & tra_pFP11_Strsp into cluster OddsH
14	noY	multiple	Gonnet	10	0.2		n	alignment poor & confusing; treat as odds
Profile								
15	2, 10	profile	Gonnet	4	0.2	1QX0	y	
16	1, 15	profile	Gonnet	4	0.2	1P4D & 1QX0	y	
17	13, 16	profile	Gonnet	4	0.2	1QX0	y	moved sequences lacking helicase domains to cluster Odds
18	12, 17	profile	Gonnet	4	0.2	1QX0	y	
19	18, OddsH	profile	Gonnet	4	0.2	1QX0	y	manual editing significant, especially N-term to MotifII poorly aligned OddsH sequences moved to OddsH2 remaining sequences truncated to relaxase domains

Table 4-2 Accession numbers and titles of relaxase sequences

sequence annotation	NCBI accession #	sequence annotation	NCBI accession #
traI F Escco	BAA97974	mobA RSF1010 Escco	CAA28520
traI p1658/97 Escco	AAO49548	repB RSF1010	AAB22064
pre pE194 Escco	AAQ98619	mobA pAB6 Neime	AAD31795
traI pAPEC-O2-R Escco	YP_190115	mobA pSJ7.4 Neigo	AAO45530
traI R100 Escco	CAA39337	mobB pDOJH10L Biflo	AAN15156
traI pSLT Salty	AAL23509	mobA pTB6 Biflo	BAD89595
traI pG8786 Yerpe	YP_093987	mobA pBLO1 Biflo	AAN31778
traI pED208 Salty	AAM90727 AAK72121	mob pNAC2 Biflo	NP_848160
traI pYJ016 Vibvu	NP_932226	mobB pKJ36 Biflo	AAG43281
traI Ch VV20663 Vibvu	AAO07605	hypo pSF118-44 Lacsa	YP_163783
traI pPBPR1 Phopr	CAG17960	traA pWCFS103 Lacpl	YP_133752
traI pLPP Legpn Paris	CAH17216	hypo pSF118-20 Lacsa	YP_163738
trwC R388 Escco	CAA44853	traA pMRC01 Lacla	AAC55993
trwC pXcB Xanci	AAO72099	traA Ch Llacc01002488 Lacla	ZP_00381644
trwC pXAC64 Xanax	NP_644759	traA pIP501 Plapl	AAA99466
traC pWW0 Psepu	NP_542915	mobL Ch ECA1659 Erwca	YP_049758
traC pDTG1 Psepu	NP_863125	traA Ch plu1828 Pholu	NP_929105
traC pBI709 Psepu	AAP57243	mobL Ch ECA2898 Erwca	YP_050989
traI pCU1 Escco	AAD27542	nes pSK41 Staau	AAC61938
traI R46 Salty	AAL13397	nes pGO1 Staau	AAB09712
trwC Ch ELI2871 Eryli	ZP_00377630	hypo Ch SC4349 Salen	YP_219336
hypo Ch ELI2182 Eryli	ZP_00376941	hypo Ch ECA1065 Erwca	YP_049172
tra Ch Saro02002574 Novar	ZP_00302710	bmgA Ch BT3143 Bacth	AAO78249
tra Ch Npun02004382 Nospu	ZP_00111530	bmgA Ch BT1126 Bacth	AAO76233
traI pCC7120gamma Nossp	NP_478459	bmgA Ch BF3285 Bacfr	CAH08980
tra Ch Npun02003762 Nospu	ZP_00110457	bmgA Ch Bacfr	AAL29920
traI pLPL Legpn Lens	CAH17351	mocA Tn4399 Bacfr	AAA98401
traI Ch lpg2077 Legpn Phil	YP_096090	hypo Ch PG0868 Porgi	AAQ66015
traA pCE2 Coref	BAC19762	bmgA Ch BT4622 Bacth	AAO79727
traA pNG2 Cordi	NP_863184	hypo pBFY46 Bacfr	YP_087130
traA pTET3 Corgl	NP_478092	mocA pBF9343 Bacfr	CAH05728
traA pGA2 Corgl	AAO18209	orf1 Ch Bacth	AAG17843
traA pCG4 Corgl	AAG00272	hypo Ch BF0132 Bacfr	YP_097415
traA pCE3 Coref	BAC19788	bmgA Ch BT0101 Bacth	AAO75208
tra pVT2 Mycav	AAL23621	hypo Ch BF1249 Bacfr	CAH06968
traA2 pAA1 Artau	AAS20144	bmgA Ch BF1367 Bacfr	YP_098652
traA Ch Nocsp	AAV52093	bmgA Ch BT2305 Bacth	AAO77412
traA pKB1 Gorwe	NP_954808	hypo Ch PG1489 Porgi	AAQ66534
traA Ch BlinB01003615 Breli	ZP_00378014	hypo Ch BF1759 Bacfr	CAH07458
traA pGOX1 Gluox	YP_190362	pnf2870 pNF2 Nocfa	YP_122139
tra pFP11 Strsp	YP_220461	virD2 pEN2701 Strsp	NP_862041
tra pFP1 Strsp	YP_220493	mobAE Ch Lacla	Q48665
traA pNAC3 Biflo	NP_848156	mobAE1 ch LLMOBAMAT Lacla	CAA61995
traA1 pSymA Sinme	NP_435751	ltrB pRS01 Lacla	AAB06502
traA2 pSymB Sinme	CAC49066	pcfG pCF10 Entfa	YP_195793
traA AT Agrtu	NP_535485	hypo pTEF2 Entfa	NP_817049
traA p42d Rhiet	AAM54881	mobA Ch gbs1121 Strag	NP_735567
traA Ch MBNC03003747 Messp	ZP_00193296	smc Ch Ssui801000289 Strsu	ZP_00333025
traA Ch lpl0169 Legpn	CAH14398	virD2 Ch Ssui801000426 Strsu	ZP_00332895
traA Ch lpp0183 Legpn	CAH11330	hypo Ch gbs1338 Strag	NP_735775
traA Ch Legpn	AAG45149	hypo Tn5252 Strag	NP_688252
traA Ch lpg1241 Legpn	YP_095272	hypo Tn5252 Strpn	NP_345530
traA pGOX3 Gluox	YP_190433	rlx pIP834 Entfa	AAF72355
traA Ch mll0964 Meslo	NP_102651	rlx Ch EF2303 Entfa	NP_815959
traA Ch AE008200.1 Agrtu	AAK88593	mobA/repB pRAS3.2 Aersa	AAK97758
traA Ch Atu4855 Agrtu	NP_535333	mobA/repB pRAS3.1 Aersa	AAK97751
traA pRi1724 Agrrh	BAB16231	mobA pTF-FC2 Acife	A43256
traA pRiA4b Agrtu	BAB47249	mobA pTC-F14 Acica	AAP04747
traA p42a Rhiet	AAO43541	nikB pCTX-M3 Citfr	NP_775011
traA pNGR234a Rhisp	T02782	mobA pEL60 Erwam	AAQ97916
traA Ti Agrtu	AAK91091	nikB pO113 Escco	AAQ17653
traA pTiC58 Agrtu	AAC17212	nikB pSERB1 Escco	AAT94234
traA pTi-SAKURA Agrtu	BAA87734	nikB R64	BAA78021
traA pHCG3 Olica	CAG28509	nikB CollB-P9 Escco	BAA75140
mobA pXF5823 Xylfa	AAK13432	nikB pSC138 Salen	AAS76381
mobA DN1 Dicno	NP_073212	mobA pDC3000A Psesy	NP_808687
mobA pIE1115 unceu	CAC05678	mobA pRA2 Pseal	AAD40339
mobA pIE1130 uncultured	CAB75594	mobA Ch PSPTO1093 Psesy	NP_790927
mobA pVM111 Pasmu	CAD55845	traI R751 Entae	NP_044272
mobA pFL190 Expve	AAV49024		

4.7 References

1. de la Cruz, F. & Davies, J. (2000) *Trends Microbiol* **8,** 128-133.

2. Burrus, V. & Waldor, M. K. (2004) *Res Microbiol* **155,** 376-386.

3. Lederberg, J. & Tatum, E. L. (1946) *Nature (London)* **158**.

4. Mazel, D. & Davies, J. (1999) *Cell Mol Life Sci* **56,** 742-754.

5. Weigel, L. M., Donlan, R. M., Shin, D. H., Jensen, B., Clark, N. C., McDougal, L. K., Zhu, W., Musser, K. A., Thompson, J., Kohlerschmidt, D*., et al.* (2007) *Antimicrobial agents and chemotherapy* **51,** 231-238.

6. Weigel, L. M., Clewell, D. B., Gill, S. R., Clark, N. C., McDougal, L. K., Flannagan, S. E., Kolonay, J. F., Shetty, J., Killgore, G. E., & Tenover, F. C. (2003) *Science (New York, N.Y* **302,** 1569-1571.

7. (2004) *Mmwr* **53,** 322-323.

8. (2002) *Mmwr* **51,** 565-567.

9. (2002) *Jama* **288,** 2116.

10. Lujan, S. & Redinbo, M. (2004) (University of North Carolina, Unpublished).

11. van Mansfeld, A. D., van Teeffelen, H. A., Baas, P. D., & Jansz, H. S. (1986) *Nucleic acids research* **14,** 4229-4238.

12. Dyda, F. & Hickman, A. B. (2003) *Structure (Camb)* **11,** 1310-1311.

13. Llosa, M., Gomis-Ruth, F. X., Coll, M., & de la Cruz Fd, F. (2002) *Molecular microbiology* **45,** 1-8.

14. Howard, M. T., Nelson, W. C., & Matson, S. W. (1995) *The Journal of biological chemistry* **270,** 28381-28386.

15. Byrd, D. R., Sampson, J. K., Ragonese, H. M., & Matson, S. W. (2002) *The Journal of biological chemistry* **277,** 42645-42653.

16. Grandoso, G., Avila, P., Cayon, A., Hernando, M. A., Llosa, M., & de la Cruz, F. (2000) *Journal of molecular biology* **295,** 1163-1172.

17. Lujan, S. A., Guogas, L. M., Ragonese, H., Matson, S. W., & Redinbo, M. R. (2007) *Proceedings of the National Academy of Sciences of the United States of America* **104,** 12282-12287.

18. Gonzalez-Perez, B., Lucas, M., Cooke, L. A., Vyle, J. S., de la Cruz, F., & Moncalian, G. (2007) *The EMBO journal* **26,** 3847-3857.

19. Guasch, A., Lucas, M., Moncalian, G., Cabezas, M., Perez-Luque, R., Gomis-Ruth, F. X., de la Cruz, F., & Coll, M. (2003) *Nat Struct Biol* **10,** 1002-1010.

20. Datta, S., Larkin, C., & Schildbach, J. F. (2003) *Structure (Camb)* **11,** 1369-1379.

21. Boer, R., Russi, S., Guasch, A., Lucas, M., Blanco, A. G., Perez-Luque, R., Coll, M., & de la Cruz, F. (2006) *Journal of molecular biology* **358,** 857-869.

22. Vezzi, A., Campanaro, S., D'Angelo, M., Simonato, F., Vitulo, N., Lauro, F., Cestaro, A., Malacrida, G., Simionati, B., Cannata, N.*, et al.* (2004) (Unpublished).

23. Sikora, B., Eoff, R. L., Matson, S. W., & Raney, K. D. (2006) *The Journal of biological chemistry* **281,** 36110-36116.

24. Lechner, R. L. & Richardson, C. C. (1983) *The Journal of biological chemistry* **258,** 11185-11196.

25. Waters, V. L. & Guiney, D. G. (1993) *Molecular microbiology* **9,** 1123-1130.

26. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997) *Nucleic acids research* **25,** 3389-3402.

27. Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994) *Nucleic acids research* **22,** 4673-4680.

28. Hall, T. A. (1999) in *Nucl. Acids. Symp. Ser.*, pp. 95-98.

29. Felsenstein, J. (1989) *Cladistics* **5,** 164-166.

30. Page, R. D. M. (1996) *Computer Applications in the Biosciences* **12,** 357-358.

5 Appendix 1: Attempted Real-Time Relaxase Kinetics

Equilibrium binding and single-time point activity data has been reported for F TraI (1-5) and for R388 TrwC (6-8), in the presence of various substrates, metal ions, mutations, and buffers . However, no steady-state kinetic assays may be found in the literature, leaving a hole in our understanding of relaxase mechanisms. Such studies are hampered by very tight substrate and product binding (1), inefficiencies associated with discontinuous assays, and difficulties and safety issues of radiolabeling. In addition, the formation of covalent phosphotyrosine intermediates with the enzyme prompted previous investigators to use high enzyme concentrations prohibitive of steady state measurements. In order to unravel the details of the TraI relaxase mechanism and to better understand reported inhibition thereof (9) I undertook the design of a continuous fluorescent kinetic TraI assay. This was ultimately unsuccessful, stymied by many of the same problems associated with previous TraI activity assays.

5.1 Results and Discussion

5.1.1 Attempted continuous kinetic assays

A continuous fluorescent assay was designed to measure the kinetics of TraI mediated ssDNA cleavage and relegation. Continuous kinetic assays for TraI relaxase activity were plagued by artifacts, many of which closely resembled the expected timecourse curves. The resemblance was great enough to drive a year-long effort to make the assay work. (**Fig. 1**) Protocol variations included different reaction buffers, temperatures, enzyme and substrate concentrations, and substrate sequences. All substrates were tested to ensure that they were within desired melting temperature bounds, but the forward and reverse melting curves differed more and more markedly as the rate of temperature change increased. This hysteresis, time-dependant non-two-state behavior, may indicate either multiple folding states or simply slow melting. (**Fig. 2**) The time for maximum melting was on the order of seconds to minutes, such that the melting rate of the substrate appears to have prevented accurate kinetic measurements. The observed first order increase in fluorescence during kinetic time courses was likely do to slow equilibration, rather than to substrate cleavage, as confirmed later by negative controls. In future, continuous TraI kinetic assays should avoid reliance on hydrolysis for product formation and DNA hybridization for product detection. I have recommended that others in the lab try fluorescence anisotropy for relaxase kinetic measurements, relying on the change in oligonucleotide size for readout of cleavage and relegation.

5.2 Materials and Methods

5.2.1 Oligonucleotide design

All oligonucleotides were ordered from Integrated DNA Technologies (IDT). Substrate oligonucleotides were large loops (20-30 bases) with complementary termini (stemloops), 5' end-labeled with a fluorescein derivative (6FAM) and 3' end-labeled with a quencher (Iowa Black FQ). The internal loop sequence was based on the previously determined tightest binding F-plasmid *oriT* segment containing the TraI nick site (1). The termini were chosen from all possible complementary sequences of seven bases or less. This list was generated in Microsoft Excel and then all melting temperatures estimated with the various oligonucleotide software packages on the Rensselaer bioinformatics web server (10). Oligonucleotides were chosen for homodimer and internal pair melting temperatures <27 °C. Stem loop (5' to 3' terminal complementarity) melting temperatures were limited to >44 °C for intact oligonucleotide and <27 °C for cleaved or nicked oligonucleotides (stem oligos). A subset of these were also chosen with stem melting temperatures >61 °C and an inverted repeat on the 5' terminus capable of hairpin formation with melting temperatures between 47 and 52 °C (snapstem oligos).

5.2.2 Oligonucleotide substrate validation

Selected oligonucleotides and their simulated (synthesized) cleavage products were melted in and re-annealed in a Fluorolog spectrometer and their fluorescein fluorescence measured. The three snapstem oligonucleotides

(TACGCCTTGGGTGT*GGTGCTTTTGGTGGTGGCGTA,
ACGAGCTTGGGTGT*GGTGCTTTTGGTGGTGCTCGT,
CGTAAGTTGGGTGT*GGTGCTTTTGGTGGTCTTACG; asterisk marks nick site) and five stem oligonucleotides
(CGTAAGTTGGGTGT*GGTGCTTTTGGTGGTCTTACG,
CGTAAGTTTGCGTGGGGTGT*GGTGCTTTCTTACG,
CGTAAGTTTTCGTGGGGTGT*GGTGCTTTCTTACG,
CCTGTGTTTTTTCGTGGGGTGT*GGTGCTTTTCAGG,
CCGCCTTGTTTTTTCGTGGGGTGT*GGTGCTTGCGG) with experimental melting temperatures within the above bounds were used in kinetic assays.

5.2.3 Attempted Kinetic Assays

Assay buffers were used as previously described (9). Briefly, kinetic assays were performed in a reaction buffer of 150 mM NaCl, 50 mM Tris-HCl pH 7.5, 6 mM $MgCl_2$ and 20% glycerol. Substrate oligonucleotide concentrations varied from 100-3600 nM. TraI relaxase domain (N300) concentrations varied from 0.5-10 nM. Reaction components (sans enzyme) were mixed at 37 °C. Pre-reaction fluorescence readings were taken to represent 0 time. Reactions were initiated by protein addition and rapid tituration, and immediately (in <10 s) read by in a BMG Pherastar microwell plate spectrophotometer (plate reader). Reactions were followed until apparent equilibria were reached.

5.2.4 Data Analysis

Nonlinear regression was used to analyze timcourses. Fluorescence timecourse data was fitted to first order exponentials. The regression parameters were then used to determine the initial reaction rates. Initial rates were plotted versus substrate concentration to construct apparent Michaelis-Menton velocity curves. These were then fitted with the simple Michaelis-Menton equation and subjected to Cornish-Bowden/Eisenthal direct linear plot analyses (11-13), both in order to determine apparent maximum reaction velocities ($V_{max.app}$) and Michalis constants ($K_{m.app}$).

5.3 Figure Legends

Figure 1. Probable equilibration artifact.

F TraI DNA “cleavage” curve displaying an increase in fluorescence originally attributed to active ssDNA cleavage. This curve fit well to a first order exponential rise to maximum, as expected for simple Michaelis-Menton enzyme kinetics, and was used to get the first estimated TraI cleavage rate. Unfortunately, curves such as this were eventually shown to be artifactual, likely due to an equilibration effect after addition of protein to the reaction mixture.

Figure 2. Hysteretic melting curves.

DNA melting (F) and refolding (R) for an example hairpin substrate oligonucleotide (designated oligo 7). The gap between the melting curves is due to hysteresis, i.e. in this case, slow or multi-state folding. This likely accounts for both the appearance of the curve in Figure 1, and the lack of real kinetic data.

Figure 5-1 Probable equilibration artifact

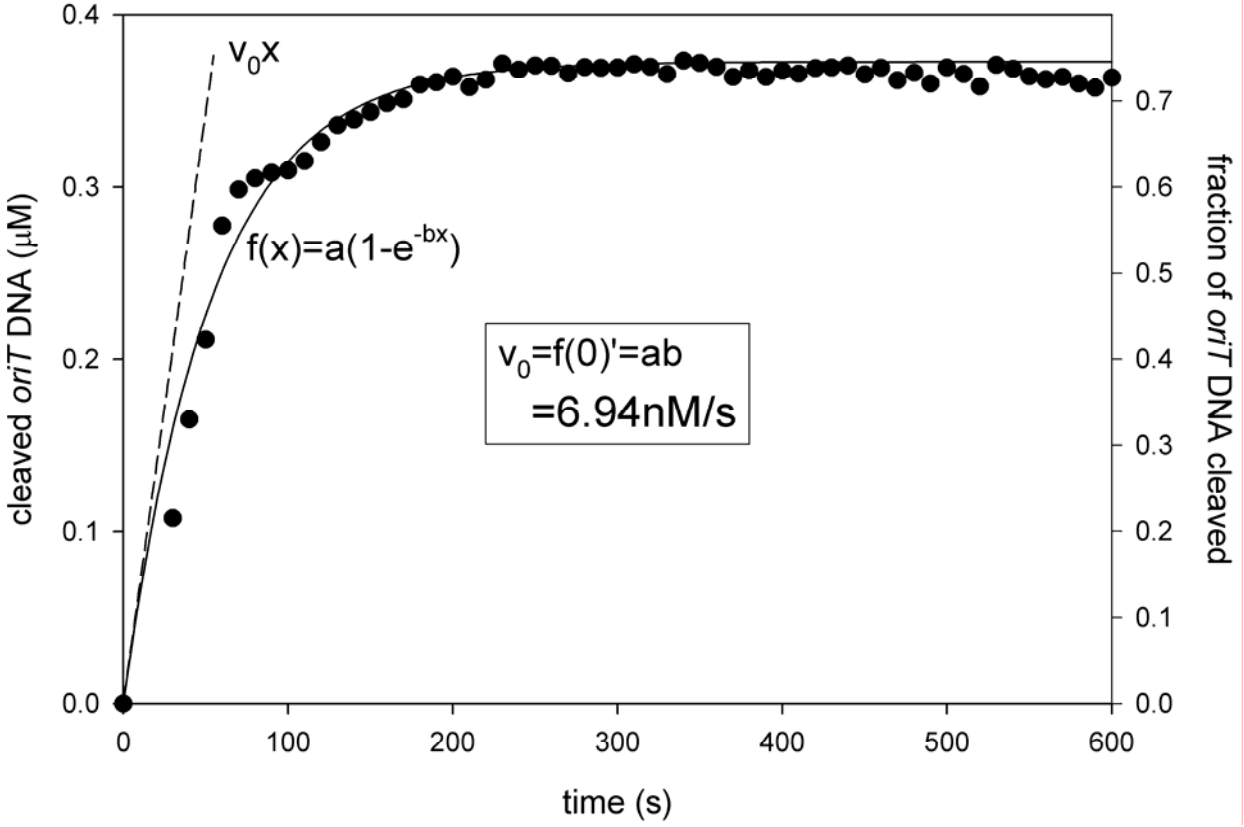
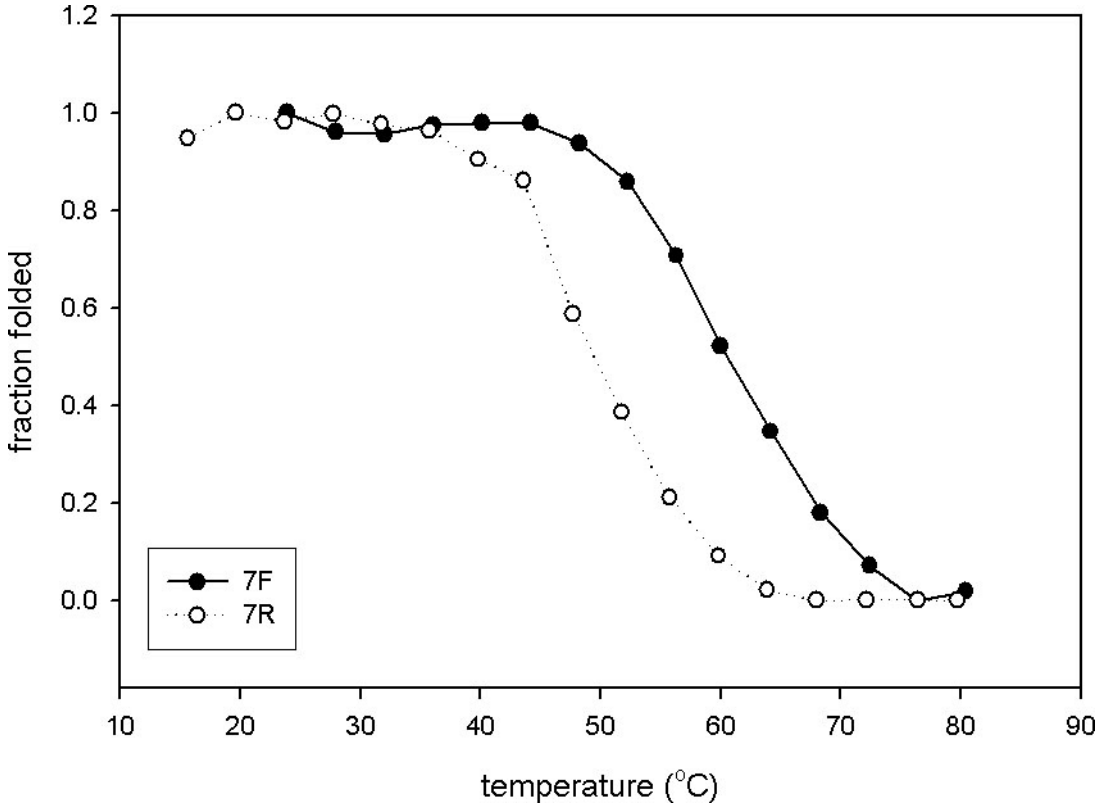


Figure 5-2 Hysteretic melting curves



5.4 References

1. Stern, J. C. & Schildbach, J. F. (2001) *Biochemistry* **40,** 11586-11595.

2. Larkin, C., Datta, S., Harley, M. J., Anderson, B. J., Ebie, A., Hargreaves, V., & Schildbach, J. F. (2005) *Structure (Camb)* **13,** 1533-1544.

3. Byrd, D. R. & Matson, S. W. (1997) *Mol Microbiol* **25,** 1011-1022.

4. Matson, S. W., Sampson, J. K., & Byrd, D. R. (2001) *The Journal of biological chemistry* **276,** 2372-2379.

5. Byrd, D. R., Sampson, J. K., Ragonese, H. M., & Matson, S. W. (2002) *The Journal of biological chemistry* **277,** 42645-42653.

6. Gonzalez-Perez, B., Lucas, M., Cooke, L. A., Vyle, J. S., de la Cruz, F., & Moncalian, G. (2007) *The EMBO journal* **26,** 3847-3857.

7. Guasch, A., Lucas, M., Moncalian, G., Cabezas, M., Perez-Luque, R., Gomis-Ruth, F. X., de la Cruz, F., & Coll, M. (2003) *Nat Struct Biol* **10,** 1002-1010.

8. Grandoso, G., Avila, P., Cayon, A., Hernando, M. A., Llosa, M., & de la Cruz, F. (2000) *J Mol Biol* **295,** 1163-1172.

9. Lujan, S. A., Guogas, L. M., Ragonese, H., Matson, S. W., & Redinbo, M. R. (2007) *Proceedings of the National Academy of Sciences of the United States of America* **104,** 12282-12287.

10. Zuker, M. (2003) *Nucleic acids research* **31,** 3406-3415.

11. Eisenthal, R. & Cornish-Bowden, A. (1974) *Biochem J* **139,** 715-720.

12. Cornish-Bowden, A. & Eisenthal, R. (1974) *Biochem J* **139,** 721-730.

13. Cornish-Bowden, A. & Eisenthal, R. (1978) *Biochim Biophys Acta* **523,** 268-272.

6 Appendix 2: Progress Toward Further F Plasmid Protein Structures

6.1 TraI constructs

3D-PSSM tertiary structure prediction was used to analyze all F TraI regions. The known relaxase and helicase regions were correctly identified as such, and a novel domain was predicted, from residues 300-590, with homology to Ffh and FtsY bacterial GTPase domains (1). These predictions were confirmed to within 95% confidence with the Tertiary Structure Prediction Metaserver, using the 3D-Jury method for correlation results from the best benchmarked tertiary structure prediction services (2, 3). We used limited proteolysis on full length TraI to determine the boundaries of globular domains. Mass spectrometric analysis of trypsin digest fragments was inconclusive, but did yield enough information to infer domain boundary through assumption of parsimony. Five globular domains were indicated, corresponding to the protease resistant relaxase fragment (residues ~1-235), two novel domains of unknown character (~235-600, roughly corresponding to the putative GTPase, and ~600-900), and two subdomains of the helicase region (~900-1100 and ~1100-1500), in addition to many smaller

fragments in the carboxy-terminal 250 residues. Tertiary structure prediction failed for the third globular domain and the C-terminus. TraI constructs 309-590 (putative GTPase) and 1-590 (N590) were expressed with the Impact expression system, purified and crystallized. Irreproducible crystal hits were observed for both constructs, but crystallization attempts were abandoned as priorities shifted to the relaxase and helicase domains. 309-590 crystals were GDP dependant. Limited proteolysis of the putative GTPase was performed in the presence and absence of GDP. The proteolysis footprint changed dramatically. The Matson lab showed that the putative GTPase will selectively bind GTP over ATP, but found no evidence for nucleotide triphosphate hydrolysis in any case. Fluorescence anisotropy-based DNA binding assays found evidence for cooperative DNA binding by both novel domains, the helicase domain, the relaxase domain, and at least a portion of the C-terminal regions. These data were collected in our lab by Jenny Xue, under Dr. Laura Guogas, and will be included in a manuscript by those investigators. The last hundred or so C-terminal residues were predicted to adopt a coiled-coil structure by Coils (4), Paircoil2 (5), and Multicoil (6). Since the Matson lab discovered an interaction between the C-terminal and TraM (7, 8), Dr. Gougas has solved the structure of the first, most conserved, C-terminal domain, sans the predicted coiled coil. The TraI regions between the relaxase and the helicase and the far C-terminus still lack structural characterization.

6.2 The F relaxosome

Interactions have been postulated between the TraI C-terminal region and TraM (based upon gel shifts and TraM dependent nicking stimulation (7)),

between the putative TraI GTPase region and TraD (in analogy with the Ffh-FtsY pair (9, 10)), between TraM and TraD (based on a variety of data in the pKM101, R1, and F systems (11-15)), and between various TraI domains (based on the DNA binding data from Xue and Guogas, above). The multitude of interactions makes it likely that the relaxosome may be a more compact entity that could be expected from its very high molecular weight and large number of subunits. If this is the case, then crystal structures of the relaxosome and/or of its various component complexes should be resolvable. The Matson lab had proven capable of reconstituting active relaxosomes in vitro, in low concentrations. The first step in structural characterization of the full relaxosome may be small angle x-ray scattering experiments using reconstituted relaxosomes in solution. Small angle scattering should reveal the shape of the overall envelope of the structure. Enough structural data and homology models exist for each component that the envelope shape may be sufficient for the modeling of the relaxosome, which could lead up to new or improved interaction hypotheses leading to the full high resolution crystal structure. The Redinbo and Matson labs have accomplished the over expression and crystallography-scale purification of each individual relaxosome component. I will begin these studies in the spring of 2008, thought they will most likely have to be completed by others after my tenure at UNC comes to a close.

6.3 References

1. Kelley, L. A., MacCallum, R. M., & Sternberg, M. J. (2000) *Journal of molecular biology* **299,** 499-520.

2. Ginalski, K. & Rychlewski, L. (2003) *Nucleic acids research* **31,** 3291-3292.

3. Ginalski, K., Elofsson, A., Fischer, D., & Rychlewski, L. (2003) *Bioinformatics (Oxford, England)* **19,** 1015-1018.

4. Lupas, A., Van Dyke, M., & Stock, J. (1991) *Science (New York, N.Y* **252,** 1162-1164.

5. McDonnell, A. V., Jiang, T., Keating, A. E., & Berger, B. (2006) *Bioinformatics (Oxford, England)* **22,** 356-358.

6. Wolf, E., Kim, P. S., & Berger, B. (1997) *Protein Sci* **6,** 1179-1189.

7. Ragonese, H., Haisch, D., Villareal, E., Choi, J. H., & Matson, S. W. (2007) *Molecular microbiology* **63,** 1173-1184.

8. Matson, S. W. & Ragonese, H. (2005) *Journal of bacteriology* **187,** 697-706.

9. Ramirez, U. D., Minasov, G., Focia, P. J., Stroud, R. M., Walter, P., Kuhn, P., & Freymann, D. M. (2002) *Journal of molecular biology* **320,** 783-799.

10. Focia, P. J., Shepotinovskaya, I. V., Seidler, J. A., & Freymann, D. M. (2004) *Science (New York, N.Y* **303,** 373-377.

11. Lu, J., Edwards, R. A., Wong, J. J., Manchak, J., Scott, P. G., Frost, L. S., & Glover, J. N. (2006) *The EMBO journal* **25,** 2930-2939.

12. Lu, J. & Frost, L. S. (2005) *Journal of bacteriology* **187,** 4767-4773.

13. Beranek, A., Zettl, M., Lorenzoni, K., Schauer, A., Manhart, M., & Koraimann, G. (2004) *Journal of bacteriology* **186,** 6999-7006.

14. Schmidt-Eisenlohr, H., Domke, N., & Baron, C. (1999) *Journal of bacteriology* **181,** 5563-5571.

15. Disque-Kochem, C. & Dreiseikelmann, B. (1997) *Journal of bacteriology* **179,** 6133-6137.

7 Appendix 3: Progress Toward a Comprehensive Real-Time Mating Assay

7.1 Previous attempts

Most current mating assays rely on endpoint analysis, as opposed to being continuous readouts. This makes it difficult to assay both inhibitor concentrations and exposure times, the two fundamental metrics of antibiotic and pharmaceutical efficacy. In 2005, de la Cruz et al. demonstrated a real-time mating assay using a fluorescent reporter (*lux* gene) inserted into an expression plasmid, downstream of a T7 promoter and an artificial R388 oriT sequence (1). When donor cells containing an oriT deficient R388 plasmid and the plasmid with the fluorescent protein gene were mated to a DE3 cell line, the fluorescent protein was expressed. The level of fluorescence at any time point was proportional to the number of transconjugation events. It was stated that this assay would allow a high-throughput screen of potential inhibitors. While clever, this assay relied on several assumptions, chief among them that donor and recipient concentrations remained unchanged for the duration of the experiment. This assumption places limitations on the viability of the assay for inhibitor

testing, namely that inhibitors that kill donor or recipient cells will be misrepresented. It could be argued that any initial high-throughput hit thus gleaned could be further characterized by more complete assays after the fact. This means that a real-time assay capable of complete characterization of all cell lines would present economies of scale and time. In addition, in the existing assay, the transferred plasmid is itself non-conjugative, which does not accurately model the exponential nature of plasmid propagation. Thus any replacement assay should place the transconjugant marker on the transferred plasmid itself.

7.2 Three fluors and complete characterization

We believe that the solution to complete characterization lies in the unique fluorescent marking of each cell line. This idea has gone through several incarnations, of which I will describe only the most recent. This involves three plasmids, each with a fluorescent marker unique to each cell line (donors, recipients, and transconjugants).

The current incarnation is based on the pCU1 plasmid system. Rebekah Potts has inserted a Venus yellow fluorescent protein gene into the T7 operon of a pET28a plasmid. She used PCR to pull out the entire operon, destroy the *lac* operator, and add Spe1 restriction sites to either end. The pCU1 plasmid has a unique Spe1 site outside of any known operons. The T7-Venus operon will be inserted into purified pCU1. This pCU1 variant (transconjugant marker) will be

electroporated into the intended donor cell line, HB101. The resulting cells will be made chemically competent in preparation for transformation with a pET11c variant encoding a second fluorescent protein (mCFP; the donor marker) and a ampicillin resistance gene. This plasmid should have low levels of constitutive fluorescent protein expression inside donor cells, analogous to that seen with the pET11c plasmid in JS11 cells during previous transfer complementation experiments. The recipient cells will be standard BL21 (DE3) cells bearing a pET28a plasmid (kanamycin resistance) encoding a third fluorescent protein (mCherry).

This assay will use a 96-well format and with three fluorescent readouts which can be read in a microtiter plate reader such as a BMG Pherastar. Donor cells will fluoresce cyan, recipients red, and transconjugants both red and yellow. Cell counts may be calculated from standard curves relating cell count to net fluorescence at each wavelength. Initial calibration of standard curves will utilize the antibiotic resistances unique to each cell line using either agar plates colon counts or the Most Probable Number method of cell counting using dilution turbidity. The ultimate goal would be to establish an optimal mixture of cells and media that could be frozen in 96-well plates, ready for use in mix-and-measure type high-throughput assays.

7.3 References

1. Fernandez-Lopez, R., Machon, C., Longshaw, C. M., Martin, S., Molin, S., Zechner, E. L., Espinosa, M., Lanka, E., & de la Cruz, F. (2005) *Microbiology (Reading, England)* **151,** 3517-3526.