THE ROLES OF STRUCTURAL VARIABILITY AND AMPHIPHILICITY OF TMC278/RILPIVIRINE IN MECHANISMS OF HIV DRUG RESISTANCE AVOIDANCE AND ENHANCED ORAL BIOAVAILABILITY

By

YULIA FRENKEL

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

and

The Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Biochemistry

written under the direction of

EDDY ARNOLD

and approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey

May 2009

ABSTRACT OF THE DISSERTATION

The Roles of Structural Variability and Amphiphilicity of TMC278/Rilpivirine in Mechanisms of HIV Drug Resistance Avoidance and Enhanced Oral Bioavailability

by YULIA FRENKEL

Dissertation Director:

Eddy Arnold

Non-nucleoside reverse transcriptase inhibitors (NNRTIs) are small hydrophobic drug molecules that bind to a hydrophobic pocket of the HIV-1 reverse transcriptase (RT). A number of molecules from diarylpyrimidine (DAPY) class of NNRTIs have exceptional inhibitory properties against a broad range of known drug-resistant viral strains of HIV-1 and good oral bioavailability in humans. Biophysical, structural, and computational methods were applied to study DAPY NNRTIs with differing RT inhibitory and bioavailability properties with a goal to define the mechanisms by which the dual success of the selected DAPY NNRTIs is achieved.

Exceptional bioavailability properties of selected DAPY NNRTIs were hypothesized to stem from their ability to form self-formulated spherical drug aggregates, ~60 nm in diameter, that are trafficked from the gastrointestinal tract into systemic circulation through a lymphatic uptake pathway. Using dynamic light scattering (DLS)

and electron microscopy (EM), formation of drug aggregates was demonstrated. Drug aggregation was determined to be surfactant independent. Based on this observation, it was hypothesized that the drug molecules were able to achieve surface-active properties. Using X-ray crystallography and all atom MD simulations of the drug aggregate/water interface it was determined that amphiphilic properties of the selected DAPY NNRTI molecules depend on the availability of the protonation site at the pyrimidine ring and structural variability of the drug molecules at the linker moieties.

Ability of the late phase DAPY NNRTI drug molecules to inhibit the broad range of drug-resistant HIV-1 mutants has been previously linked to the ability of the molecules to sample multiple binding modes within the binding pocket of HIV-1 RT (Das *et al.*, 2004). Using computational methods, conformational sampling by TMC278 (a DAPY NNRTI in multiple Phase III trials) in the pocket of the wild-type and the mutant NNRTI-binding pocket (NNIBP) was evaluated. Results of this work support the concept that structural variability is important in overcoming drug resistance mechanisms.

We propose that the implementation of drug aggregation testing and the evaluation of torsional flexibility of highly hydrophobic compounds in the context of the lead identification strategies will allow for better selection of potent and orally bioavailable hydrophobic drug candidates.

Acknowledgements

I would like to express my deepest gratitude to my graduate advisor, Dr. Eddy Arnold. Eddy has been a great source of inspiration, drive, and knowledge through out my graduate years. Eddy's remarkable intuition and insight into the fields of molecular biology, chemistry, and drug development have been instrumental for the development and progress of my research. Owing to Eddy's unwavering belief that anything is possible, during the years of my studies I learned a great wealth of multidisciplinary methods and gained a multidimensional perspective on the intricacies of drug development. Besides professional mentorship, I am also thankful to Eddy for his compassionate and sympathetic nature. These past years had been filled with many experiences, some happy and some not, but regardless of the circumstance, Eddy has always been supportive and understanding. I am very thankful to him for these displays of kindness.

I would like to thank my thesis committee members: Dr. Emilio Gallicchio, Dr. Ronald M. Levy, Dr. Ann Stock, and Dr. Wilma Olson. Dr. Emilio Gallicchio provided me with tremendous help in both experimental work on molecular dynamics and in writing of paper manuscripts and the doctoral thesis. It has been a great pleasure to work with such a knowledgeable, skillful, and creative person. I greatly appreciate all of the time and effort that has been contributed by him towards creation of a very interesting and fruitful computational project. I am thankful to Dr. Levy for his help in focusing my efforts towards creating a cohesive thesis in the last few years of my graduate studies. To Dr. Stock and Dr. Olson I am thankful for helpful discussions on both experimental and professional matters.

This graduate experience would not have been as growth stimulating as it has without the help of the members of the Arnold Lab and the CABM family (past and present). I thank Dr. Deena Oren for mentoring me in the experimental design, techniques, analysis, as well as in writing and presentation skills. I am grateful to Deena for her thought provoking discussions, advice, and her optimistic outlook on life. I would like to thank Dr. Kalyan Das for providing me with valuable training, guidance, and advice. I am thankful to Arthur Clark, Jr., Dr. Thomas Mariano, and Chhaya Dharia for both tremendous technical help and for training in protein purification, crystallization, and tissue culture. I thank Dr. Steve Tuske, Dr. Stefan Sarafianos, Dr. Daniel Himmel, Tom Eck, and Dr. Gail Ferstandig Arnold for their help, support and knowledge. In CABM community, I would like to express my deepest gratitude to Madeline Frances for keeping me enrolled in graduate school, as well as for her constant support and helpful guidance in professional and personal matters. I thank Dr. Tanya Borsuk, Dr. Joseph Bauman, and Denise Elssasser for being good friends and supporters.

At last, I would like to thank my family and friends. I thank my parents for being the driving force and the stable backing to me and my family through these years. I thank Genia and Leo Frenkel for all their help. I thank friends from DaWoods for their love and friendship. Above all, I thank my best friend and supporter, my husband, Vladimir. Without his love and encouragement, I would have never been able to weather my years in graduate school.

Dedication

To my husband, Vladimir,

and

my son, Daniel

Contents

List of abbreviations

AIDS	acquired immunodeficiency syndrome
AUC	area under the curve
BCS	biopharmaceutics classification system
Caco-2	human caucasian colon adenocarcinoma epithelial cell line
DAPY	diarylpyrimidine
DATA	diaryltriazine
DLS	dynamic light scattering
DPD	dissipative particle dynamics
EC_{50}	50% effective concentration
GI	gastrointestinal
HAART	highly active antiretroviral therapy
HIV-1	human immunodeficiency virus type 1
HTS	high throughput screen
IAM	immobilized artificial membrane
kD	kilodalton
MALT	mucosa-associated lymphoid tissue
MW	molecular weight
MD	molecular dynamics
NCE	new chemical entity
NNIBP	NNRTI-binding pocket
NNRTI	non-nucleoside RT inhibitor
NRTI	nucleoside/nucleotide RT inhibitor
PCA	principal component analysis
PDB	Protein Data Bank
PE	potential energy
REMD	replica exchange molecular dynamics
RT	reverse transcriptase
TEM	transmission electron microscopy
SASA	solvent accessible surface area

SPC	simple point charge
SEDDS	self-emulsifying drug delivery system
SESA	solvent excluded surface area
SMEDDS	self-microemulsifying drug delivery system

List of tables

List of illustrations

Introduction

A biophysical study exploring aggregation properties of diarylpyrimidine (DAPY) non-nucleoside reverse transcriptase inhibitors (NNRTIs) in solutions mimicking gastric conditions of a fasting individual is reported in chapter one. Based on results of the solution studies, a correlation between ability of the drug molecules to form ~100 nm diameter spherical nanoparticles and good oral bioavailability is reported. The second chapter is dedicated to analysis of five small molecule DAPY NNRTI X-ray structures. The structural analysis reports evidence and detailed aspects of conformational variability and amphiphilic nature for a subset of studied NNRTIs. The third chapter reports a computational study on DAPY NNRTI aggregate formation. Molecular dynamics (MD) simulation results show that amphiphilic properties of DAPY NNRTIs under acidic conditions and high conformational variability play essential roles in water/aggregate interface formation for microemulsion-like drug aggregate structures. The fourth chapter investigates mechanisms used by TMC278, the highly promising DAPY NNRTI drug candidate, to overcome the effects of resistance mutations in the binding pockets of wild-type and mutant RTs. MD simulation results provide evidence of conformational variability in both drug and binding pocket residues within the time-span of the simulation.

Chapter 1. Aggregation of DAPY NNRTIs in simple solutions and relevance to oral bioavailability

Synopsis

We have examined physicochemical properties of highly hydrophobic NNRTI/DAPY compounds and explored a possible correlation with oral bioavailability. In simple solutions designed to mimic the GI environment of a fasting individual, all of the tested NNRTI compounds formed either colloidal solutions or precipitated material. As determined through DLS and EM, colloid formation was dependent on drug concentration, and pH of solution. The ability to form homogenous populations of colloidal aggregates with hydrodynamic radii <100 nm under conditions of high drug concentration and acidic pH correlated with good oral bioavailability. Based on our observations, we have proposed a model in which uptake of NNRTI/DAPY compounds into systemic circulation depends on their ability to present high concentrations of appropriately sized colloidal drug aggregates at different locations within the GI tract.

Introduction

Acquired immunodeficiency syndrome (AIDS)

AIDS is caused by the human immunodeficiency virus (HIV). HIV impairs normal activities of the immune system and therefore creates susceptibility to a broad range of life-threatening symptoms and infections (Weiss, 1993). According to the most recent report by the World Health Organization (WHO), 33 million people were infected with HIV by the end of 2007, of which 2.7 million were newly infected individuals (UNAIDS/WHO, 2008). No cure or vaccine is available for treatment or prevention of HIV/AIDS. The only currently recommended preemptive action against HIV infection is to avoid exposure to the virus; if that fails then highly active antiretroviral therapy (HAART) is offered (Hammer et al., 2008; UNAIDS/WHO, 2008). HAART consists of combinations of highly potent antiviral therapeutic agents from different antiviral classes that target multiple stages in the HIV replication cycle. The most widely used cocktails include nucleoside reverse transcriptase inhibitors (NRTIs), protease inhibitors (PIs), and NNRTIs.

Drug resistance to antiretroviral drugs is a major obstacle to long-term success in anti-HIV therapy. The HIV-1 RNA-dependent DNA polymerase, known as reverse transcriptase (RT), lacks proofreading capabilities. As a consequence, replication of the HIV-1 genome is highly error prone. The RNA template-dependent substitution error rate is 0.38 for 10 kb of HIV-1 genome per replication cycle (Boyer et al., 1992). The intrinsically high error rate, combined with a high rate of viral turnover and a very large number of virions in an infected individual, leads to rapid accumulation of mutations

(Coffin, 1995; Mansky and Temin, 1995). Under drug selection pressure, mutations of the viral targets can lead to drug resistance and loss of therapeutic efficacy for even very potent drugs. Therefore, a key focus of drug development efforts has been to design therapeutic agents that have broad cross-reactivity against all or most of the currently known drug-resistance mutations (Janssen, 2005).

RT

RT plays an essential role in the HIV-1 replication cycle. The enzyme is required for transcription of a single-stranded viral RNA genome into a double-stranded DNA form (for review of reverse transcription see Coffin *et al*., 1995). The newly created DNA can then be incorporated into the host genome. RT performs two enzymatic functions: it is a DNA polymerase and a ribonuclease H (RNase H). As a DNA polymerase, RT copies either an RNA or a DNA template into a complementary DNA sequence. As an RNase H, RT cleaves and degrades the RNA template during and after synthesis of the first DNA strand.

HIV-1 RT is an asymmetric heterodimer, with two related subunits. The p66 subunit is a 66-kDa polypeptide consisting of polymerase and RNase H domains. The polymerase domain is further subdivided into fingers, palm, thumb, and connection subdomains. The p51 subunit is a product of the same gene as p66, however, as a result of proteolytic cleavage of the p66 subunit, the C-terminal RNase H subdomain is absent in p51 (Figure 1) (Kohlstaedt et al., 1992). Although they share the same amino acid sequence, p66 and p51 subunits differ in their functions and overall structures. Catalytic

activities of DNA polymerase and RNase H are performed by the p66 subunit while as p51 serves mainly as a structural scaffold, and for binding the anti-codon stem and loop of tRNA-Lys which functions as the primer for DNA synthesis (Arts et al., 1998; Barat et al., 1989).

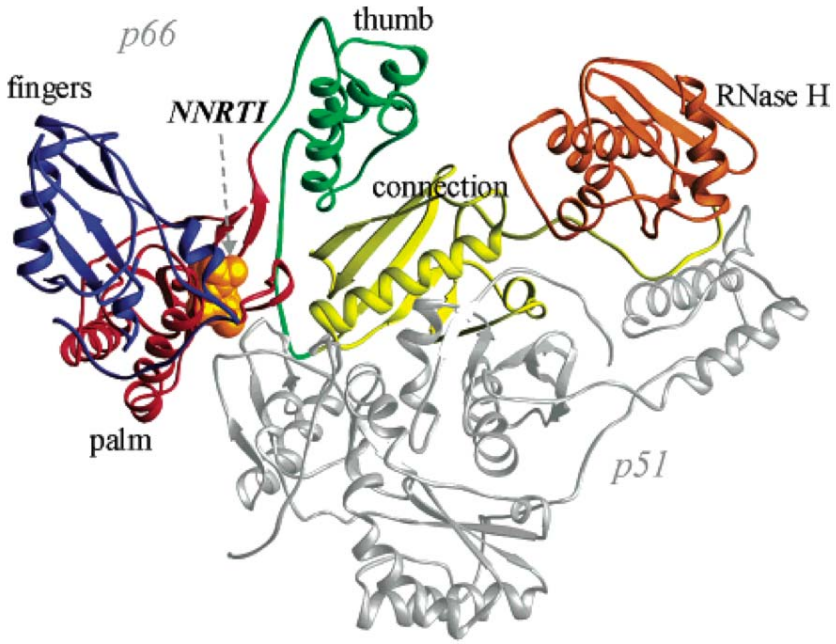


Figure 1. Ribbon diagram of HIV-1 RT p66/p51 heterodimer with space-filled atoms indicating the location of the NNRTI-binding pocket. Color assignment for the individual subdomains: p66 fingers – blue, p66 thumb – green, p66 palm – red, p66 connection – yellow, RNase H – orange, p51 – gray (adopted from Janssen et al., 2005).

NNRTIs

Due to its critical role in the HIV replication cycle, RT is a principal target for anti-HIV drug development. RT inhibitors targeting DNA polymerase function are categorized into two groups: NRTIs and NNRTIs. NRTIs are called chain terminators and are synthetic analogs of the natural dNTP substrates. NNRTIs are allosteric inhibitors that bind to a site 10 Å distant from the substrate-binding site (Das et al., 2004;

Ding et al., 1995; Kohlstaedt et al., 1992). The spatial and functional association between the two binding sites has been exploited by combination therapies where both categories of inhibitors are included to increase effectiveness of the treatment (De Clercq, 2004).

Inhibition of RT via NNRTI binding has been proposed to employ a number of possible mechanisms. One of the early models suggested that NNRTI binding caused restrictions on movements of the relative domains important for catalytic transitions of the enzyme, otherwise known as 'molecular arthritis' (Kohlstaedt et al., 1992). However, determination of multiple structures in complex with NNRTIs, where subdomain orientations have been shown to differ, suggested that relative movement of catalytically important domains is still feasible even in the presence of NNRTI. Another inhibition mechanism proposes that NNRTI binding resulted in repositioning of the three-stranded β-sheet of p66 that contained key catalytic residues Asp110, Asp185, and Asp186, therefore, forcing RT into a trapped, inactive conformation (Esnouf et al., 1995). Another mechanism suggested by Das et al., proposed that the significant repositioning of the "primer grip" that accompanies NNRTI binding may interfere with catalysis owing to inappropriate positioning of the primer terminus for catalysis (Das et al., 1996). A more detailed evaluation based on crystallographic structures suggested a distortion of key catalytic residue orientations in such fashion that binding of divalent cations Mn^{2+}/Mg^{2+} was inhibited and was proposed to weaken ability of the enzyme to bind dNTPs (Das et al., 2007). A contrary result was obtained in kinetic studies of NNRTI binding to RT where association between the subunits was not affected by the inhibitor binding, and suggested that metal coordination was not involved in the inhibitory mechanism. Though

a large number of plausible mechanisms had been proposed, presence of conflicting information continues to put under question our understanding of the NNRTI inhibitory mechanisms, therefore leaving the need for further investigation.

Presently there are four NNRTIs approved by the FDA: nevirapine /Viramune™, delavirdine/Rescriptor™, and efavirenz/Sustiva™/Stocrin™, and TMC125/etravirine/ Intelence™. NNRTIs including nevirapine and efavirenz are recognized as agents to use in first line of therapy for HIV positive patients (Figure 2A-C). However, the benefit of potency and a long half-life of these three drugs have been offset by a rapid evolution

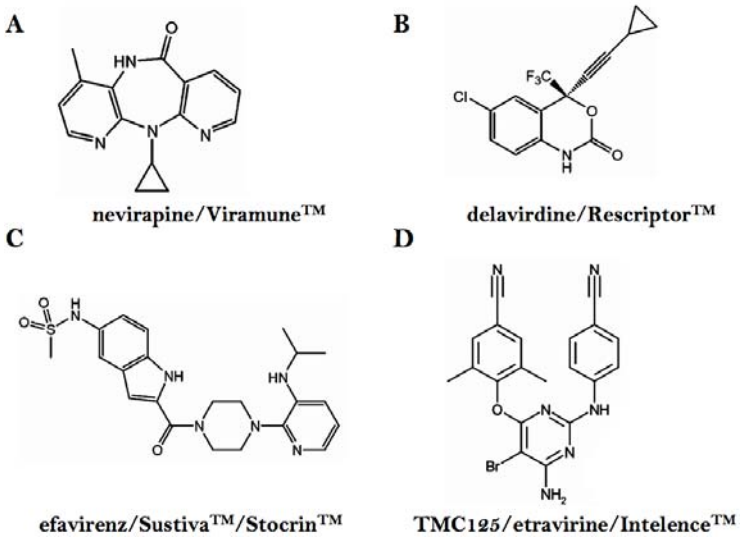


Figure 2. The chemical structures of four NNRTIs currently approved as anti-HIV/AIDS therapeutic agents.

of drug resistance. In an effort to tackle the drug resistance challenge for targeting HIV-1 RT, a multi-disciplinary structure-based effort (Das et al., 2004) targeting identification of chemical entities that can inhibit a wide range of currently known drug-resistant HIV-1 viral strains has led to the discovery of the diaryltriazine (DATA)(Ludovici et al., 2001a) and the DAPY (Figure 5 and Table 1)(Ludovici et al., 2001b) classes of NNRTIs (Gruzdev et al., 2003; Sankatsing et al., 2003). Of the DAPY NNRTIs, the FDA

approved recently TMC125 (Figure 2D) for treatment of HIV-1 infection. Concurrently, another DAPY NNRTI, TMC278/rilpivirine, has been designated as a fast tracked agent by the FDA in two Phase III clinical trials (Figure 5, Table 1).

Compared with earlier generations of NNRTIs, DAPY compounds are able to reduce viral loads of currently known drug-resistant strains with exceptional robustness. TMC278 had shown exceptional resilience towards single and double mutants with ability to inhibit 94% of mutant isolates at EC_{50} ≤10 nM, while as efavirenz and nevirapine at the same concentrations displayed only a 69% and an 18% inhibitory ability respectively (Janssen et al., 2005). Due to unprecedented drug-resistant mutant inhibition properties of DAPY NNRTIs, TMC125 and TMC278 are being considered for an updated HAART regimen for anti-retroviral-experienced HIV positive patients (Hill et al., 2007).

Bioavailability

Drug-exposed individuals are high susceptible to developing drug resistance, and require strict compliance to the prescribed regimen in order to keep disease under control. Compliance depends on personal and social factors. Statistics reveal that although significant progress has been made in improving availability of affordable therapeutic agents, compliance is still critically problematic in low- and middle-income populations worldwide (UNAIDS/WHO, 2008). Lifestyle changes associated with long-term adherence to the multi-dose treatment requires diligence and organization that are not easily practiced and thus add a significant burden on HIV positive patients. Therefore,

another central goal in treatment development against HIV/AIDS is implementation and development of therapeutic agents with high levels of bioavailability, retention, and low rate of elimination, thus ensuring extended activity of single doses and lower number of necessary doses per unit of time.

Bioavailability is a primary pharmacokinetic parameter in determining the potential success of a drug (Ahr et al., 2000). Oral bioavailability is defined as the rate and extent of drug molecule absorption into systemic circulation from the gastrointestinal (GI) tract and is governed by the drug and dosage form (Martinez and Amidon, 2002; Rosenberg et al., 1993a; Rosenberg et al., 1993b). NNRTIs are highly hydrophobic (ClogP ~5), small (300 to 500 Da) molecules that have poor solubility in water. For compounds with low solubility, oral bioavailability is of great concern. Good oral bioavailability depends strongly on dissolution and absorption of a drug into systemic circulation. The Biopharmaceutics Classification System (BCS) is a widely accepted tool to classify drug candidates based on their solubility/dissolution and intestinal permeability properties (Amidon et al., 1995). Figure 3 lists BCS classes and corresponding solubility and permeability requirements. In the earlier years of HTS implementation for drug lead identification, compounds in BCS classes I type were most preferred for drug development (Ahr et al., 2000). BCS was used to select most ideal candidates out of the large pool of potential lead compounds. However, in the recent decade the number of leads belonging to either poor solubility or poor permeability classes has greatly increased. Current statistics report that 35-40% of new chemical entities (NCEs) are known to have poor solubility properties under physiologically relevant conditions and belong to BCS class II and IV types of molecules with poor

solubility and high or low permeability properties (Amidon et al., 1995; Stegemann et al., 2007). Approximately 40% of drug candidates tested in late-stage clinical trials fail due to low efficacy for reasons including poor bioavailability that can be caused by ineffective intestinal absorption and/or undesirable metabolic activity (Li, 2001, 2003; Martinez and Amidon, 2002). Reliable bioavailability evaluations generally involve testing in both

Biopharmaceutics class	Solubility	Intestinal permeability
I	HS	HP
II	LS	HP
III	HS	LP
IV	LS	LP

Figure 3. List of criteria for the four BCS classes of compounds. HS: high solubility; largest dose dissolved in ≤ 250 ml of water over pH range 1 – 8. HP: high permeability; extent of absorption is ≥ 80%. (The figure is adopted from Amidon, et al., 1999.)

animal models and in human subjects, requiring considerable time and expense (White, 2000). *In vivo* systems for bioavailability estimation include monitoring of drug concentrations following intraduodenal dosing in animal models (usually rat and dog) and intestinal perfusion studies in both animals and human subjects (Rosenberg et al., 1993). *In vitro* test formats include: 1) Caco-2 cell monolayers for measurement of compound diffusion rate (Caldwell et al., 1998; Fung et al., 2003; Larger et al., 2002); 2) intestinal tissue for permeation measurements (Junginger et al., 1999; Yu et al., 1996); and 3) IAM chromatography (Pidgeon et al., 1995). A number of high-throughput *in vitro* methods have been designed, however, they are still several orders of magnitude

slower than desired and have other limitations as well (Ansede and Thakker, 2004; Caldwell et al., 1998; Fung et al., 2003; Li, 2003; Ungell et al., 1998; White, 2000).

Absorption of lipophilic particles into systemic circulation occurs in a number of ways with some passing through enterocytes via paracellular and/or transcellular routes, and others through the microvilli (M) cells in Peyer's patches of mucosa-associated lymphoid tissue (MALT) in intestinal lumen via transcytosis (Chen and Langer, 1998; Hodges et al., 1995; Jani et al., 1989; Kraehenbuhl and Neutra, 2000). Size exclusion (Desai et al., 1996; Eldridge et al., 1989), dissolution (Dressman and Reppas, 2000), transfer (Kostewicz et al., 2002), chylomicron (Charman and Stella, 1986; Charman et al., 1992), and surface dependent (Ikomi et al., 1999) models for absorption of lipophilic drugs address the multidimensionality of this issue, but the lack of universality of these models suggests the need for further investigation in this field.

In order to improve intestinal solubility and dissolution of Class II and IV compounds, a myriad of formulation designs have been explored. Formulation methods include salt complexation, prodrug formation, particle size reduction, self-emulsifying drug delivery system (SEDDS), self-microemulsifying drug delivery system (SMEDDS), micellar/surfactant system, and solid dispersion (Narang et al., 2007; Porter et al., 2008; Stegemann et al., 2007; Vasconcelos et al., 2007). Most of the formulation designs are focused on generation of stable, nanosize delivery units obtained either by altering the physical/chemical properties of the drug molecules or by addition of co-solvents, surfactants, and/or lipid suspensions. Unfortunately, besides the intrinsic complexity and drug specificity involved in application of some of the listed methods there are still issues with drug precipitation, low drug loading, instability, excipient-related toxicity, and first-

pass metabolism. To address these issues further investigation and improvement is necessary (Kipp, 2004; Stegemann et al., 2007).

Though many of the DAPY NNRTIs suffer from the same bioavailability problem as other BCS Class II and IV compounds, some have surprisingly good oral bioavailability properties despite their high hydrophobicity and low water solubility (Janssen et al., 2005; Lewi et al., 2004). One such example is TMC278. Unlike the recently approved drug, etravirine/TMC125, that requires granulolayered tablet formulation in order to reduce the pill size from 800 mg to 200 mg administered twice a day (Levin, 2005), TMC278 can be effectively given as a single 25 - 150 mg dose per day (Levin, 2008). Significant reduction in pill burden is anticipated to improve compliance with the new regimens. To understand the source of unique behavior of selected NNRTIs is of great importance not only to the development of orally bioavailable drugs as anti-AIDS treatments but also for other BCS Class II and IV NCE compounds circulating in lead identification pipelines.

NNRTI absorption hypothesis by Dr. Paul Janssen

After working for many years with about 4000 NNRTIs, Dr. Paul Janssen and colleagues at the Janssen Center for Molecular Design formulated a hypothesis and proposed a possible model of oral absorption for highly hydrophobic/lipophilic compounds. Even with the most potent NNRTIs there seemed to be a threshold minimum EC_{50} value of ~0.1 nM in a standard cell-based assay (Pauwels et al., 1988). Based on this observation, Dr. Janssen hypothesized that NNRTIs were absorbed as aggregates as opposed to individual molecules. Calculating the number of molecules per

cell from the EC_{50} threshold value and the number of cells in the cell culture experiments, it could be estimated that the aggregates would contain ~2 x 10^6 NNRTI molecules (with an average molecular volume of 0.38 nm^3 for a typical NNRTI) and would have a diameter of ~56 nm. Assuming that the volume and the mass of different NNRTI molecules vary, the particles were predicted to form aggregates with diameters ranging from approximately 50 to 70 nm. Concurrent studies conducted in dog models showed that high concentrations of orally administered NNRTIs were present in lymph tissue during the early phase of absorption (up to ~4 h) followed by equilibration (after ~6 h) between lymph and plasma concentrations. Building on these findings, the aggregation hypothesis stated earlier was extended to a lymphatic absorption of aggregates model in which orally administered NNRTIs may form aggregates in the GI tract where they can be absorbed by M cells of Peyer's patches of MALT and then drained into lymphatic circulation and emptied into the systemic compartment (Lewi et al., 2004). If the aggregation hypothesis is correct, it will explain not only the presence of drug concentration minima in tissue culture but also the possible mechanisms for drug delivery and uptake of other highly hydrophobic drug candidates. Therefore, to verify this hypothesis of drug aggregation and to study further behavior of NNRTI molecules in aqueous environment, we have designed a number of biophysical and computational studies where aggregate formation was monitored, characterized, and evaluated as a function of drug structure and aqueous conditions.

Results

In an effort to identify properties and mechanisms responsible for enhanced oral bioavailability of selected NNRTIs we designed a number of biophysical experiments

testing Janssen's hypothesis of aggregate formation. Using DLS and TEM we were able to verify aggregate formation for all DAPY NNRTIs available for analysis. Aggregate size and morphology correlated with oral bioavailability data for NNRTI compounds, therefore supporting the aggregate uptake model proposed by Dr. Janssen.

Aggregation behavior of NNRTIs was examined as a function of concentration, pH, stabilizing agent, co-solvent, and ionic strength using DLS and EM. A total of 36 DATA/DAPY NNRTIs were used in the DLS analysis and three compounds were evaluated by TEM. Of the 36 compounds studied, 15 had known bioavailability information (Figure 4, Table 1).

Aggregation of NNRTIs

DLS experiments were designed to detect the ability of the drug molecules to form nanoparticles, in an aqueous media that mimicked gastric conditions of a fasting individual (0.15 M HCl, pH 1.5). Drug behavior was evaluated both, in the presence and the absence, of a stabilizing agent, tyloxapol. Tyloxapol is a nonionic surfactant that forms 3.5 nm radii micelles in aqueous solutions (Regev and Zana, 1999). Initial observations are presented in Figure 5, where the behavior of Compound 15/TMC278 is observed in aqueous media, both in the presence and absence of tyloxapol. Based on the DLS percent intensity and mass plots, reported in Figure 5, in the absence of the drug the

compd	ID/generic name	AUC rat (μg·h/mL)	AUC human (μg·h/mL)	pK_a	ClogP
1	R106168	1.5		3.1	4.7
2	R120393	67.5		3.3	4.2
3	R124043	4.4		2.8	5.3
4	R126874	7.3			5.5
5	R129385	0.6		1.9	4.3
6	R132914	0.1	0.1	2.5	4.7
7	R138750	5.2	0.7	3.8	5.6
8	R147681/TMC120/dapivirine	19.4	4.7	5.8	6.3
9	R152649	8.4		2.6	5.5
10	R152929	49.3	8.2	5.1	5.3
11	R156204	6.8	0.9	3.6	7
12	R157753	0.1		2.7	5.2
13	R165335/TMC125/etravirine	0.6	0.4	3.5	5.2
14	R185545	0.6		2.0	4.1
15	R278474/TMC278/rilpivirine	9.8		5.6	5.8

Table 1. Absorption and ClogP data for 15 NNRTIs. Rat and human absorption (exposure, AUC) values were determined following oral administration of 40 mg/kg and 100 mg/kg doses, respectively (Lewi et al., 2004).

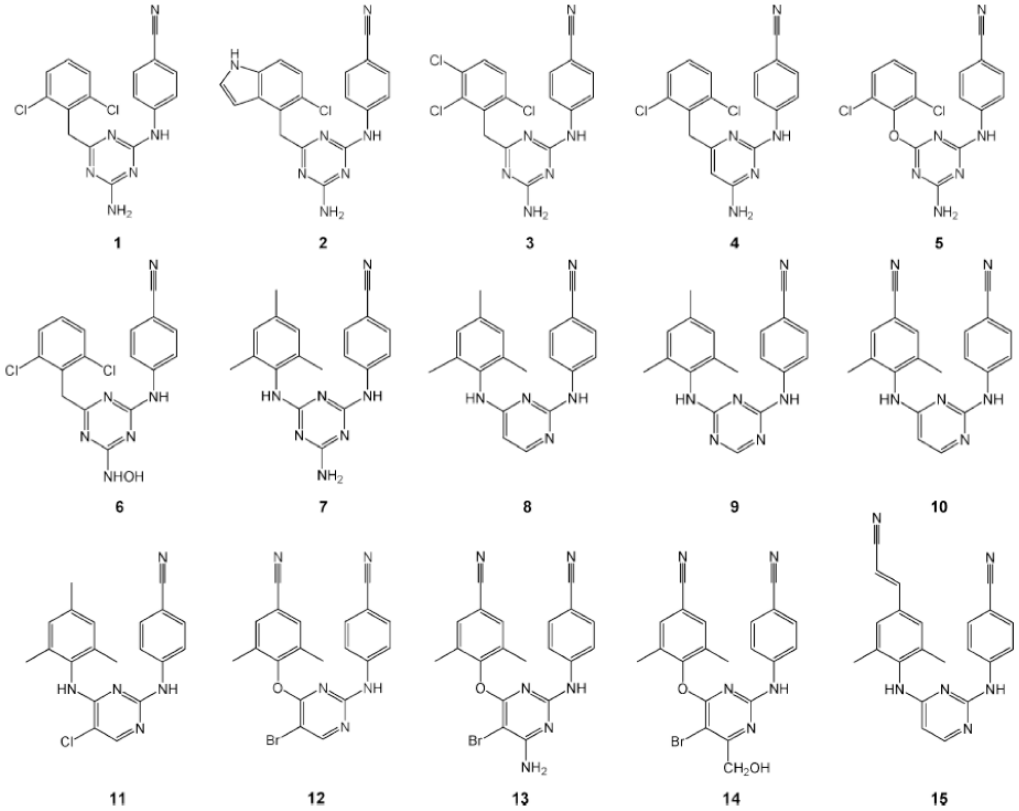


Figure 4. Chemical structures of 15 NNRTIs. Numbers correspond to compound IDs listed in Table 1.

only observed aggregated structures were those corresponding to surfactant micelles (3.5 nm). In the presence of the drug, an additional aggregate population was detected comprising much larger radii (~60 nm). To determine whether observed behavior is shared by other DATA/DAPY NNRTIs, a larger subset of drug molecules was tested in a similar fashion. All 36 NNRTIs tested via DLS showed signs of aggregation. The aggregate size varied for different compounds. As shown in Figure 6, the measured aggregate sizes fell into two distinct groups. One group of compounds had aggregate radii ranging from 30 to 110 nm, while the measured radii in the other group were greater than 250 nm.

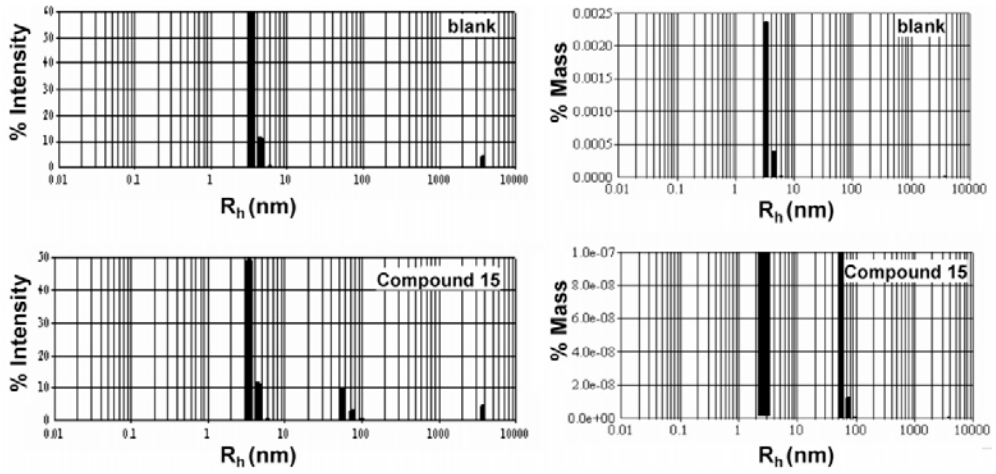


Figure 5. DLS% intensity and% mass versus R_h histograms for solution without drug (blank) and the same containing 0.1 mM Compound 15/TMC278. The blank contains 0.1% tyloxapol in 0.15 M HCl, pH 1.5. The aggregate populations at 3.5 nm radii correspond to tyloxapol micelles. The peaks at ~3000 nm correspond to instrumental artifacts due to a small amount of background scattering.

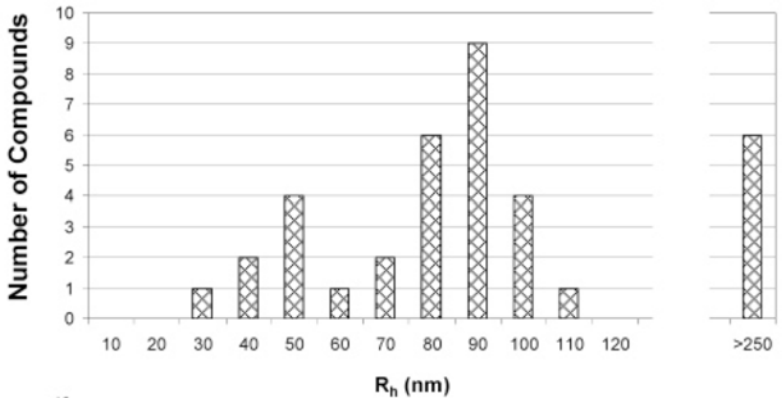


Figure 6. Size distribution of aggregates expressed as number of compounds forming particles with mean hydrodynamic radii (R_h) as determined using DLS for 36 NNRTIs at 0.1 mM concentration in simple stomach-mimicking solutions (0.15 M HCl, pH 1.5, with and without stabilizing agent tyloxapol). NNRTI aggregate sizes fell into two groups: 1) radii ranging from 30 to 110 nm; and 2) radii greater than or equal to 250 nm.

Compounds in the small aggregate size groups formed homogenous aggregate populations. Aggregate population homogeneity was assessed based on the transparency of the final experimental solutions and stability in acquisition of DLS measurements. The large aggregate size group was best characterized as precipitate based on the cloudiness of the final solutions, settling at the bottom, and high polydispersity (>40%). We have analyzed compounds from both size groups in terms of their morphology, sensitivity to solutions conditions (variation on pH, surfactant, co-solvent), and correlation with oral bioavailability information from animal models and humans.

Morphology of NNRTI aggregates

Assessment of aggregate morphology is important the understanding of possible formation, stabilization, and cellular uptake mechanisms of drug nanoparticles. Several aggregate morphologies have been observed in small molecule colloidal formation including spherical, lamellar, and fibrous, to name a few (Iavicoli et al., 2009). Dr. Janssen's model assumes that drug aggregate morphology is approximately spherical. Analogously, DLS size and molecular weight estimate is dependent on the shape model. To verify Janssen's hypothesis and the accuracy of DLS-based predictions, we analyzed TMC120, TMC125, and TMC278 using negative staining TEM (see Figure 7). As shown in Figure 7, our control (blank) comprised of HCl and detergent only, is in agreement with DLS results reported in Figure 5. Formation of aggregated structures is observed only in samples containing drug molecules. For compounds TMC120 and

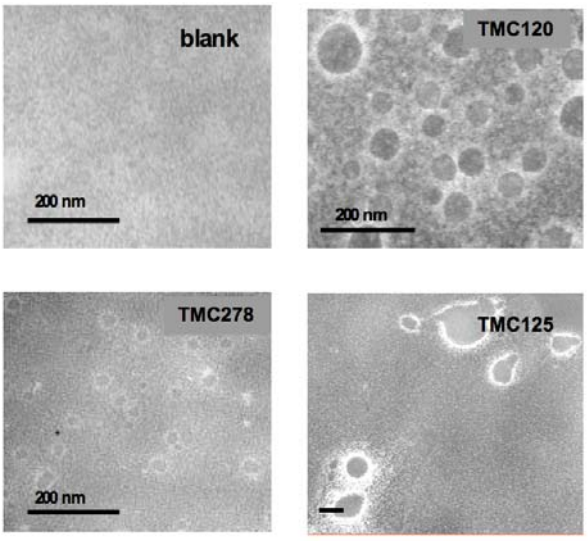


Figure 7. TEM visualization of blank, and 0.25 mM TMC120, TMC278, and TMC125 in 0.1% tyloxapol in 0.15 M HCl, pH 1.5. The blank contains 0.1% tyloxapol in 0.15 M HCl, pH 1.5. The black bar corresponds to 200 nm.

TMC278, formation of spherical aggregates is clearly visualized. TMC120 structures are generally larger in size with 40 – 60 nm in diameter, while TMC278 aggregate size ranges from 20 to 40 nm. In contrast to TMC120 and TMC278, TMC125 aggregate structures were not evenly distributed and varied substantially in their size and morphology. Colloidal structures of TMC125, shown in Figure 7, are large in size (>200 nm in diameter) and irregular in shape. Based on these findings, we conclude that some of the NNRTIs (such as TMC120 and TMC278) can form spherical particles while others form non-homogeneous populations of large, irregular structures, possibly best described as precipitate.

Detergent and co-solvent sensitivity of NNRTI drug aggregation

1) Detergent sensitivity of NNRTI aggregation

DLS evaluation of NNRTI aggregation was performed in aqueous solutions in the presence and absence of surface-active agents. Surfactants were initially introduced into the system in an attempt to stabilize the DLS measurements, which were erratic presumably because of rapid changes in the sizes of the drug aggregates. A series of non-ionic detergents including β-octylglucoside, Triton X-100, and tyloxapol were tested for their effect on the DLS measurements. Tyloxapol was found to have the most favorable properties in terms of stability of data collection together with forming relatively uniform, small detergent micelles within an easily distinguishable and narrow hydrodynamic size range.

We compared aggregation tendencies of 15 NNRTI compounds in the presence

and absence of tyloxapol in solutions simulating gastric conditions of a fasting individual. As depicted in Figure 8, DLS measurements of 15 NNRTIs at pH 1.5 in the absence and presence of tyloxapol all showed the presence of either colloidal aggregated structures (R_h <110 nm) or precipitate material (R_h >250 nm). In the presence of the detergent, a ~20 nm increase in the average hydrodynamic radii of the aggregated species were observed. One of the 15 tested NNRTIs, R157753 exhibited exceptional detergent dependent aggregation behavior where precipitated material formed in the absence of surfactant while homogeneous aggregate populations were detected in its presence. A small subset of drug molecules was unable to form stable aggregate populations either in the absence

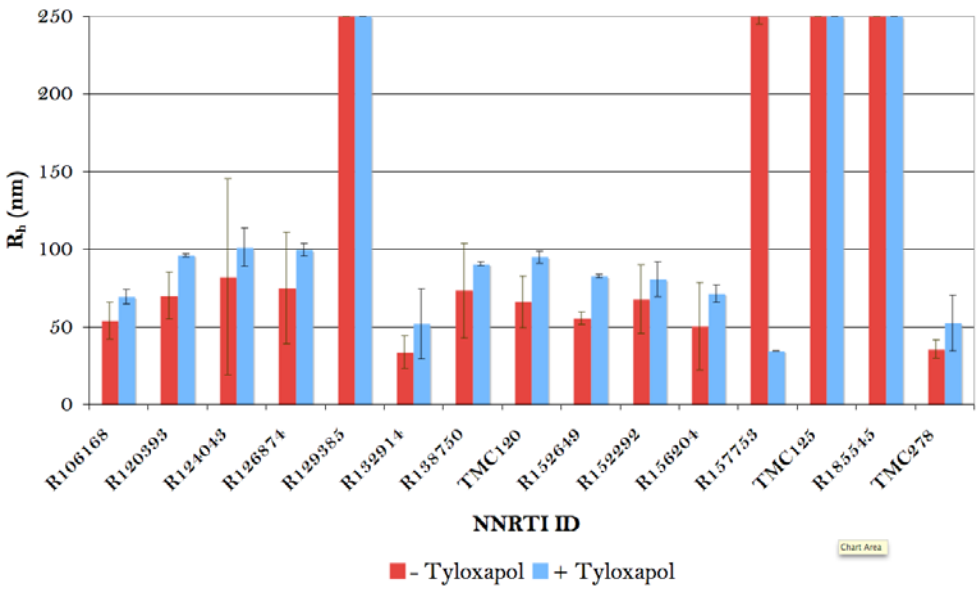


Figure 8. Average R_h measurements for 15 DAPY NNRTIs in the presence and absence of tyloxapol. Color assignment: red: DLS measurements for 0.1 mM drugs in 0.15 M HCl, 0.1% tyloxapol; blue: DLS measurements for 0.1 mM drugs in 0.15 M HCl. Error bars reflect the standard deviations calculated based on four independent measurements for each inhibitor (see details in Materials and Methods).

or in the presence of the detergent: they were R129385, TMC125, and R185545. Based on the fact that all other solution components were kept constant, these observations suggest that aggregate formation is primarily a function of drug molecules' atomic composition.

Drug transition through the GI tract is accompanied by changes in pH, therefore in addition to evaluating behavior of NNRTIs under gastric conditions that simulate

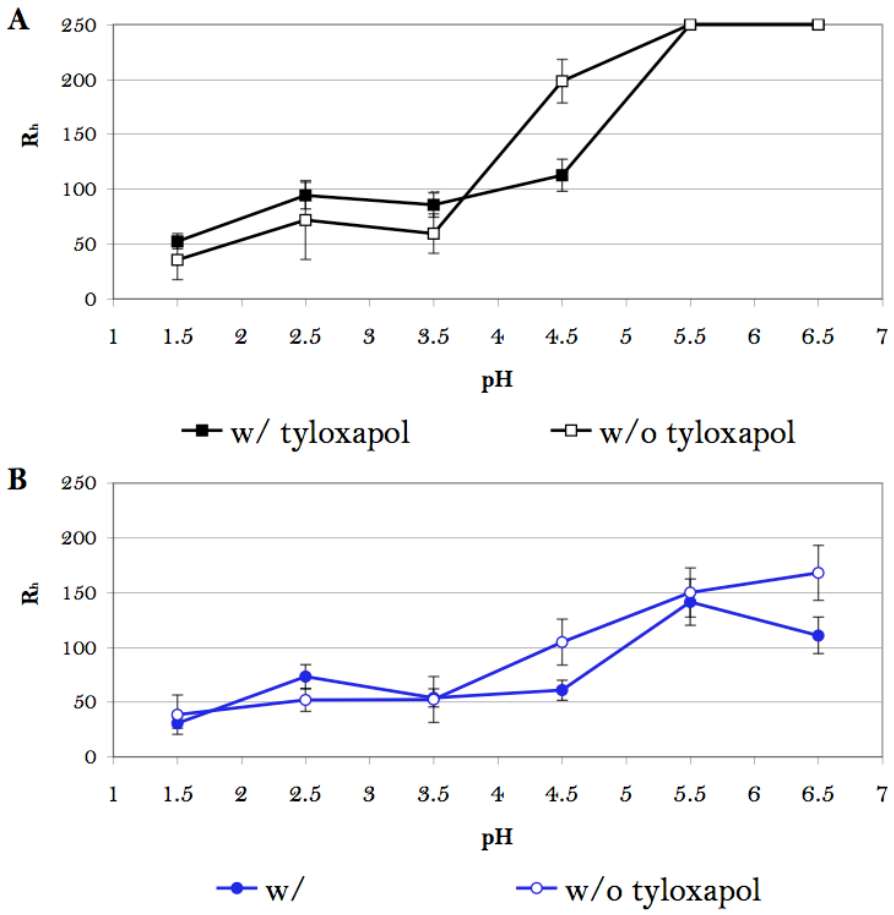


Figure 9. TMC278 co-solvent and surfactant dependent aggregation at different pHs. DLS measured hydrodynamic radius distribution as a function of solution pH for 0.1 mM TMC278 pre-dissolved in either DMSO (A) or PEG400 (B) in the presence and the absence of 0.1% tyloxapol (v/v). Color assignment: black solid squares: pre-dissolved in DMSO, with tyloxapol; black clear squares: pre-dissolved in DMSO, without tyloxapol; blue solid circles - pre-dissolved in PEG400, with tyloxapol; blue clear circles - pre-dissolved in PEG400, without tyloxapol.

stomach dissolution, we also evaluated detergent dependence of aggregation at different pHs for one of the NNRTI molecules, TMC278. We monitored aggregate formation for a 0.1 mM drug in the presence and absence of tyloxapol. As shown in Figure 9A, the average size of TMC278 aggregates in aqueous solutions tended to increase with increase in pH. Our results show that at the most acidic conditions (pH 1.5), slightly larger aggregates were observed in the presence of tyloxapol than in its absence (drug aggregate radius of 52.5 ± 6.8 nm versus 35.5 ± 17.7 nm, respectively). This trend persisted in the pH range from 2.5 to 3.5, where in the presence of tyloxapol the average aggregate size was greater by ~20 nm than in the tyloxapol-free solutions. At pH 4.5, the trend observed at lower pHs was reversed and smaller aggregate size was observed for samples with tyloxapol (112.5 ± 14.5 nm) versus without (198.5 ± 20.0 nm). Formation of the homogeneous aggregate population at the higher pHs was not observed and is best characterized as a precipitate. Based on these observations we conclude that solution pH affects the size of drug aggregates in the presence and absence of detergent in similar fashion, however, the NNRTI aggregation tendencies are detergent insensitive only under acidic conditions, while as at higher pH, detergent sensitivity is more pronounced.

2) Co-solvent sensitivity of NNRTI aggregation

In addition to testing detergent sensitivity of 15 NNRTI compounds, we also evaluated their dependence on the nature of the added co-solvent. Due to their high hydrophobicity, all of the NNRTIs in the study were pre-dissolved in either DMSO or PEG400. Earlier efforts to dissolve dry drug material directly in the aqueous solutions

proved to be unsuccessful. Traditionally, biophysical studies involving hydrophobic drug molecules used either 100% DMSO or PEG400, the latter is also used in drug formulations for many hydrophobic and hydrophilic compounds. In our studies, 20 or 50 mM drug stock solutions in either of the two co-solvents were used, with final DMSO or PEG400 concentration at 0.2% (v/v). In animal studies, initial formulation designs targeting oral delivery were based on PEG400 mixtures and their concentrations are comparable. In an effort to determine whether drug aggregate formation reported earlier was dependent on the presence of excipients, we compared aggregation behavior of 15 NNRTI compounds in both co-solvents.

Comparison of drug aggregate formation in DMSO and PEG400 had shown that NNRTI particle formation is insensitive to the nature of the additive. Control experiments containing DMSO or PEG400 in the absence of drugs did not show formation of aggregated structures. In Figure 10, we depict results of co-solvent dependent aggregation at experimentally relevant drug concentrations and pHs for TMC125 and TMC278. TMC278 formed smaller aggregated structures at low concentration in PEG400 than in DMSO. Though slight differences were observed, the overall trend of aggregation was preserved and at pH 6.5 under conditions of 1 and 0.1 mM the drug precipitated while at pH 1.5 homogeneous populations of colloidal nanoparticles were formed. In the case of TMC125, neither DMSO nor PEG400 conditions showed formation of aggregated material in a systematic fashion. In PEG400, few measurements at low concentrations were possible, however, those measurements were not reproducible and were shown to reflect the difference between the two systems.

In general, both compounds formed slightly smaller and at times more stable aggregated structures in PEG400 then in DMSO (Figure 9). Though we saw slight differences in

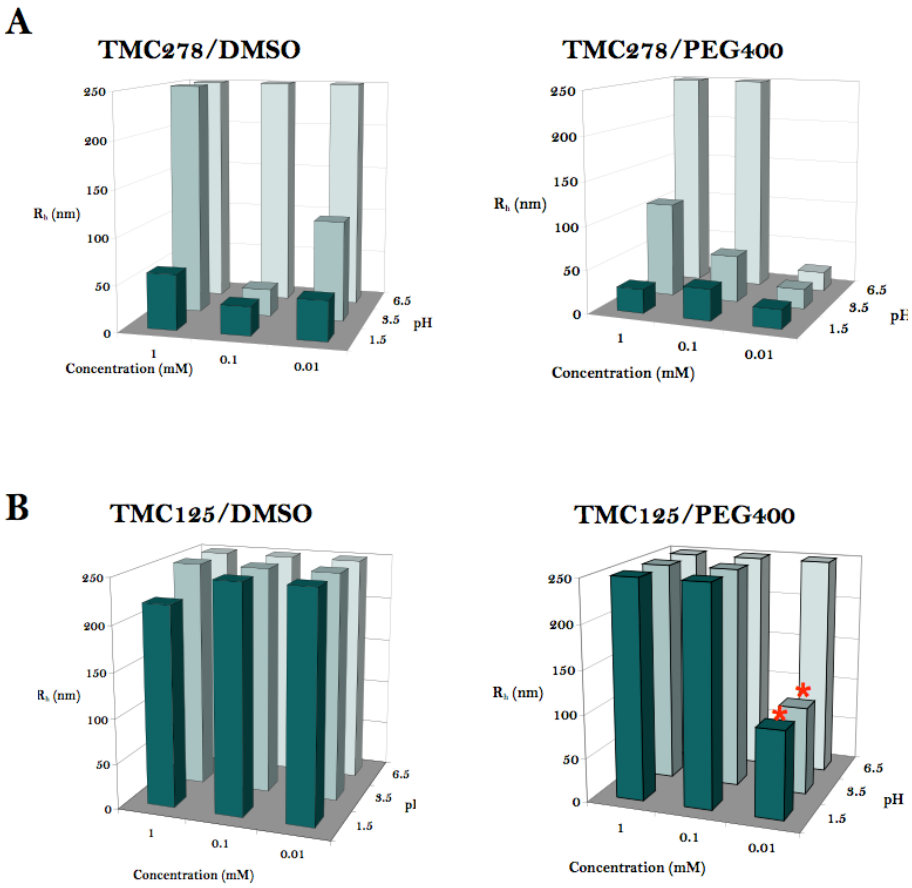


Figure 10. DLS measured R_h distribution plots for DMSO or PEG400 pre-dissolved TMC278 and TMC125 at physiologically relevant pH and drug concentration conditions. Color assignment: dark green: pH 1.5; medium green: pH 3.5; light green/blue: pH 6.5. *A small subset of measurements was used to generate the average value. None of the experimental solutions contained tyloxapol. Tiffany Tsay, an undergraduate assistant, conducted data collection.

aggregate behavior, the overall trends of aggregation were preserved for both compounds, i.e. TMC278 was able to form aggregated structures under most of the sampled conditions in pH and concentration dependent manner while TMC125 failed to do so.

3) Combined effect of a detergent and a co-solvent on NNRTI aggregation

We also evaluated the combined effect of co-solvent and surfactant presence on drug aggregate formation under physiologically relevant conditions. Figure 9 depicts results of this investigation. The aggregate size measurements reported in Figure 10 suggest that addition of co-solvent did not change the overall trend of NNRTI aggregation with different aggregate formation characteristics. In the study exploring dual effects of surfactant and co-solvent on drug aggregation, we looked at 0.1 mM TMC278 pre-dissolved in DMSO (Figure 9A) and in PEG400 (Figure 9B) in the presence and absence of tyloxapol. As discussed earlier with increase in pH, drug aggregate size had been observed to increase. The aggregate size of drug nanoparticles containing PEG400 was overall smaller than that with DMSO. At the most acidic conditions (pH 1.5), slightly larger aggregates were observed in the presence of tyloxapol (average drug aggregate radius 52.5 ± 6.8 nm) than in its absence (35.5 ± 17.7 nm) for TMC278 aggregates in DMSO. In the case of PEG400 containing pre-dissolved material, the size difference was within the standard deviation. At pH 4.5, the trend observed at lower pH values was reversed and smaller aggregate size was observed for samples with tyloxapol (113 ± 15 nm in DMSO, 53 ± 10 nm in PEG400) versus without (199 ± 20 nm in DMSO, 105 ± 20 nm in PEG400). Formation of the homogeneous aggregate population at the higher pHs was not observed for DMSO samples and was best characterized as a precipitate. At high pH, PEG400 containing solutions produced large colloidal structures that were more stable in the presence of detergent than in its absence. In summary, presence of surfactant in either DMSO or PEG400 resulted in increase in

aggregate size and stability under acidic conditions and size reduction with improved aggregate stability at higher pHs.

Aggregate stability

Stability of NNRTI drug aggregates was tested as a function of temperature and time. We tested temperature-dependent aggregation for three drug molecules that were earlier determined to have different aggregation tendencies ranging from moderate to heavy (TMC120, TMC278, and TMC125). Figure 11 shows samples of experimental

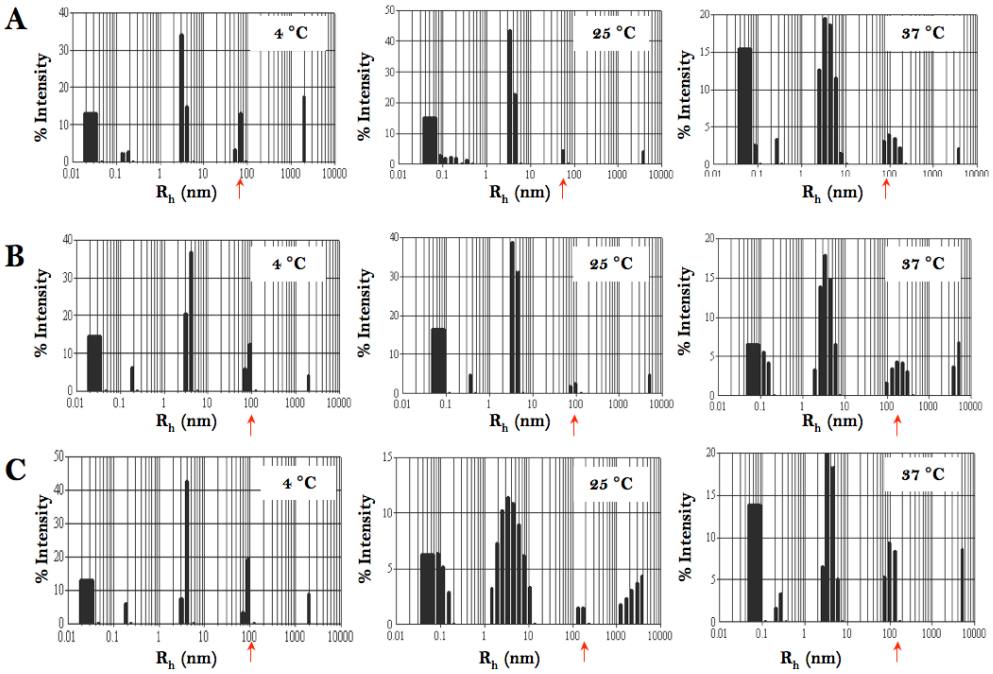


Figure 11. Light scattering percent intensity histograms for TMC120 at 4ºC, 25ºC, and 37ºC. Tested samples contained 0.1 mM (A), 0.5 mM (B), and 1.0 mM (C) concentrations of the drug in 0.1% tyloxapol in 0.15 M HCl, pH 1.5. Red arrows point to the most populated aggregate peaks based on percent mass estimates. The aggregate populations at 3.5 nm radii correspond to tyloxapol micelles. The peaks at <1 nm and ~3000 nm correspond to instrumental artifacts due to solution and background scattering, respectively.

data for TMC120 at three different concentrations and temperatures. As the data suggest, the aggregate formation was observed at all tested conditions. In general, the higher temperature conditions correlated with increase in population heterogeneity as well as a modest increase in average aggregate size (59 ± 15 nm at 25°C versus 72 ±18 nm at 37°C). Comparison of three drug molecules with different aggregation tendencies at 4ºC, 25ºC, and 37ºC is shown in Figure 12A. The histogram suggests that aggregation properties of the drug molecules are conserved in the 4°C to 37°C temperature range: that is, TMC120 and TMC278 formed homogeneous aggregate populations at all three tested temperatures while TMC125 was unable to do so. Differences in measured aggregate sizes for TMC120 and TMC278 at 25ºC and 37ºC were within one standard deviation, yet heterogeneity was more pronounced at higher temperatures. In order to test whether aggregate formation was stable at temperatures above 37°C, we also performed a temperature scan for one of the tested compounds (TMC278). As shown in Figure 12B, in the temperature range from 4°C to 85°C, aggregate formation was consistently observed throughout. Aggregate size fluctuation was observed to be within one standard deviation with exception of measurements conducted at 15ºC.

Based on these observations we concluded that NNRTI compounds that were able to form homogenous populations of drug aggregate at 25ºC were also able to do so at the physiologically relevant temperature of 37ºC. Even though the trend of drug aggregation was preserved across tested temperatures the aggregate population heterogeneity at 37ºC was higher. Therefore we concluded that 25ºC was the most optimal temperature for studying aggregation trends of NNRTIs in simple solutions using DLS.

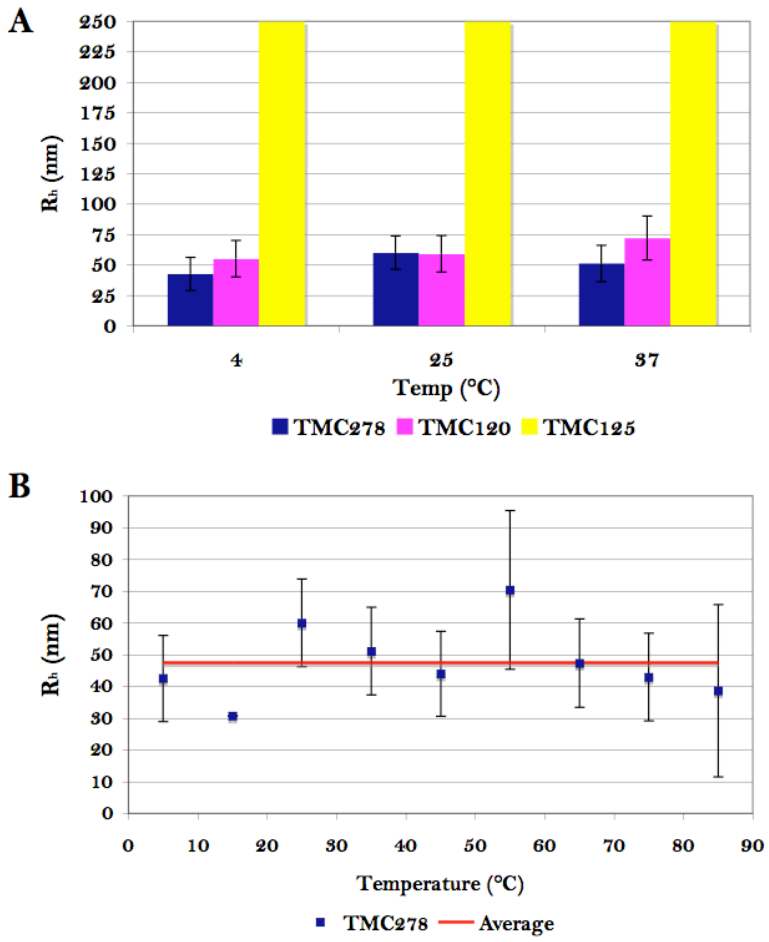


Figure 12. Average R_h versus temperature plots for TMC278, TMC120, and TMC125. A. Comparison of drug aggregation for TMC120, TMC278 and TMC125 at 4°C, 25°C, and 37°C. B. Temperature scan from 4°C to 85°C for TMC278. Tested solutions contained 0.1 mM drug in 0.1% tyloxapol, 0.15 M HCl pH 1.5. Error bars reflect the standard deviations calculated based on tree independent measurements for each inhibitor (see details in Materials and Methods). Average reflects linear regression over the average aggregate size measurements of TMC278.

Correlation between pH and concentration sensitivity and oral bioavailability

Of the 36 NNRTIs studied, 15 compounds had known animal (rat and dog) and, in six cases, human absorption data (see Table 1, Figure 4). The absorption of a drug

molecule, which is determined as the exposure or area under the curve (AUC) of the plasma concentration curve as a function of time following oral administration, has been used in this study as a correlate of oral bioavailability. The rat and human absorption data for DATA/DAPY NNRTIs were shown to correlate better than dog and human absorption data (Lewi et al., 2004). Therefore, only rat and human absorption data were used to search for correlations with DLS and EM findings. Compounds with the highest exposure in the rat were R152929, TMC120, and TMC278 (Table 1). The same three compounds were also highly bioavailable in human subjects. In addition to analyzing rat and human exposure data, Lewi *et al*. also characterized NNRTIs in terms of pKa, ClogP, and other experimental and computed physicochemical parameters (Table 1).

To further investigate the possible dependence of NNRTI aggregation on concentration and pH, we generated a dilution series from 10 mM to 0.001 mM in the physiologically relevant pH range from 1.5 to 6.5. Based on our results, all aggregate sizes for NNRTIs increased with increasing compound concentration. As an example of concentration-dependent behavior, EM images of TMC278 at 0.25 mM and 0.5 mM concentrations are presented in Figure 13. The two-fold increase in concentration led to

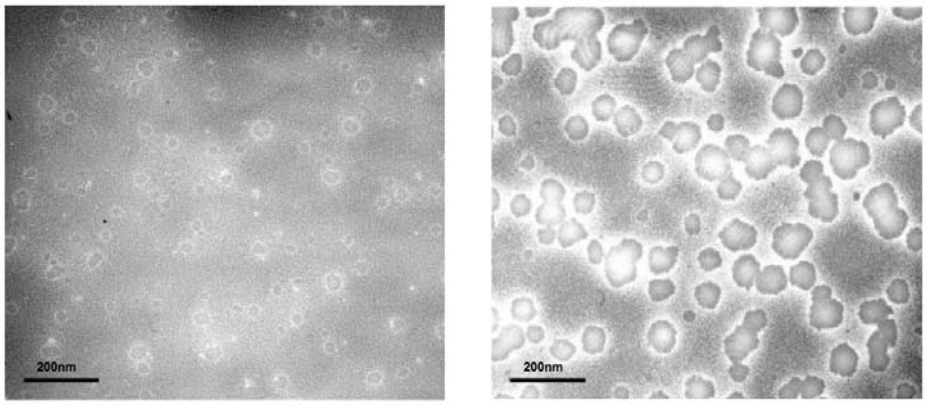


Figure 13. TEM images of TMC278 aggregates at 0.25 mM and 0.5 mM drug concentration.

an increase in aggregate radii from ~30 nm to ~60 nm. At the higher concentration aggregates coalesced into very large structures. However, even at the higher compound concentration a minority of small aggregates was present.

The effect of pH variation on the size distribution of NNRTI aggregates was evaluated for three compounds (R129385, R152929, and R156204) with different rat/human absorption (Table 1). Based on DLS analysis, the pH dependence of particle size distribution for all three compounds at 0.1 mM was reflected in an increase in aggregate size with increasing pH (Figure 14). R152929 was the most bioavailable

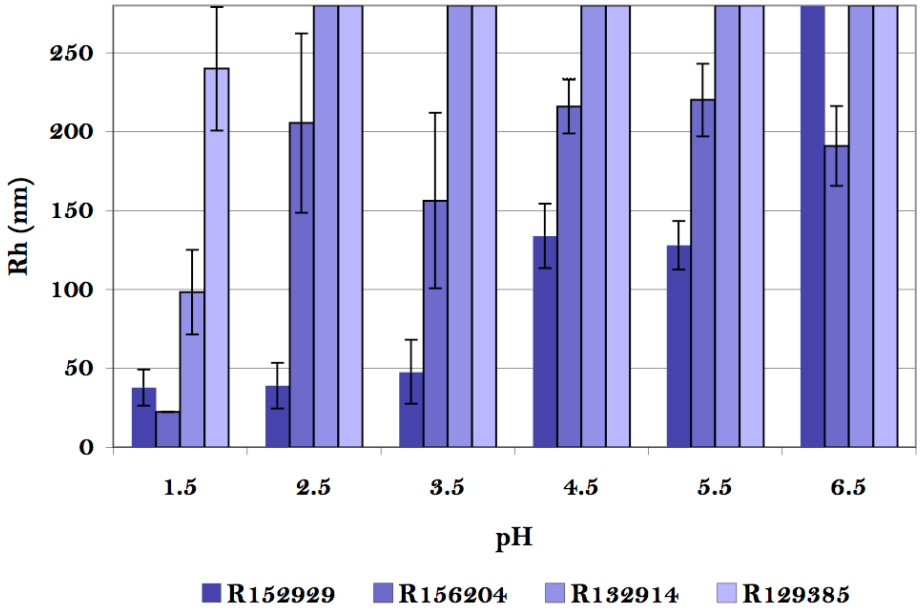


Figure 14. DLS determined drug aggregate R_h at physiologically relevant pHs for four DAPY NNRTIs with differing oral bioavailabilities. Color scale assignment: dark blue – compound with good oral absorption (AUC rat/human >5 µg·h/mL); medium blue – intermediate absorption (1 µg·h/mL <AUC in rat/human <5 µg·h/mL); light blue – poor absorption (AUC in rat/human <1 µg·h/mL).

inhibitor of the three compared and its aggregation behavior was least affected by changes in pH, with aggregate radii below 50 nm until pH reached 4.5, while R129385

formed aggregates with radii greater than 200 nm at pH values 2.5 and above. R129385 had the most pH-sensitive aggregation behavior and was the least bioavailable of the three compounds. R129385 was classified as poorly bioavailable based on oral administration studies (see Table 1); when compared to compounds with better absorption, the behavior of R129385 was distinct because it was highly susceptible to formation of large structures at pH values above 1.5 while with compounds with better absorption such as R152929 and R156204, the aggregate sizes increased more gradually with increasing pH.

Compounds with good and poor absorption show distinct aggregation behavior

Analysis of R_h measured by DLS using various gastric-mimicking conditions for 15 NNRTIs with known absorption revealed that compounds with more favorable absorption form aggregated material of smaller size than the less bioavailable molecules. As reported earlier, all of the 36 NNRTIs tested in the study formed aggregated species. However, the size distribution analysis clearly had shown the presence of two distinct types of colloidal forms. Small aggregates were determined to be spherical and their size ranging from 30 to 100 nm (see Figures A2 and A3 and Table 1) while the other group of compounds formed irregular structures with aggregate size larger than 250 nm in radii. Of the 36 compounds, 15 had known bioavailability (Table 1) and were analyzed in the context of size distribution histogram shown in Figure 6. As shown in Figure 15, although the smaller size aggregates were observed to be formed by compounds with various oral bioavailability, compounds that formed aggregates with greater than 250 nm radii invariably had poor absorption (AUC<1.0 μg·h/mL in the rat).

We evaluated feasibility of using aggregate behavior in simple solution as a classification for oral absorption and determined that this can be a plausible method for such purpose. From the data discussed earlier and depicted in Figure 15, the success of *in vitro*/*in vivo* classification of favorable versus poor absorption group based on aggregation behavior was determined to be 86.7% for 0.15 mM HCl with surfactant at pH 1.5 (13 of 15 assigned correctly to absorption group) and 93.3% for 20 mM malonic acid without surfactant at pH 1.5 (14 of 15 correctly assigned). Based on Fisher's exact

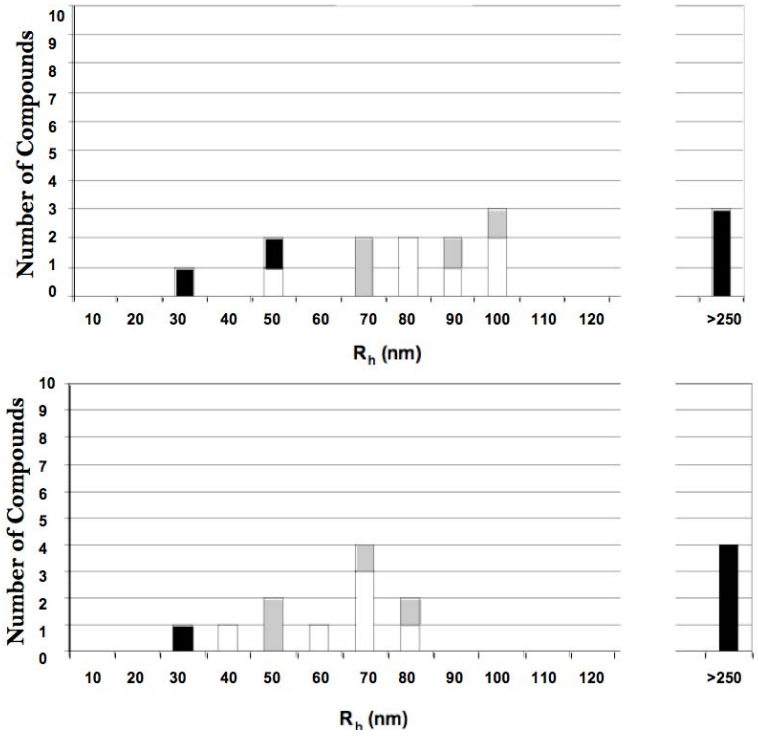


Figure 15. Size distribution of aggregates expressed as the number of compounds forming particles with mean R_h as determined using DLS for 15 NNRTIs with known bioavailability. Drug aggregation was evaluated at 0.1 mM drug concentration in simple stomach-mimicking solutions (top – 0.15 M HCl, pH 1.5; bottom – 20 mM Malonate buffer, pH 1.5). NNRTI aggregate sizes fell into two groups: 1) radii ranging from 30 to 110 nm; and 2) radii greater than or equal to 250 nm. Grayscale assignment: white – good absorption (AUC rat/human >5 μg·h/mL); gray – intermediate absorption (1 μg·h/mL <AUC in rat/human <5 μg·h/mL); black – poor absorption (AUC in rat/human <1 μg·h/mL).

test, (JMP), the two-tailed probability for the 2x2 aggregation/exposure table is 0.004. The overall prediction accuracy of absorption and probably bioavailability was better in the absence of the surfactant. Nonetheless, it has to be stressed that presence of the surfactant seemed to facilitate conversion from precipitated to aggregated states for compounds with poor absorption.

Different concentration and pH dependent aggregation behavior of compounds with good and poor absorption

After studying aggregation as a function of concentration and pH as single variables, we investigated the combinatorial effects of these parameters on aggregation of NNRTIs with different rat/human absorption properties (Figure 16). Compounds R152929, R156204, and R129385 were considered to have good, intermediate, and poor absorption, respectively, in rats and human subjects (Table 1). While all compounds formed large structures, possibly representing precipitate, at higher concentrations and solution pH values, the data indicated that compounds with good or intermediate absorption formed smaller aggregates (30 nm $<R_h<$100 nm) at higher concentrations and pH while those with poor absorption formed larger structures. R152929 was the only drug in the given series that was able to form small, stable, homogeneous populations of drug aggregates at the highest tested concentration of 10 mM at pH 1.5. Intermediate bioavailability compounds, R156204, formed more stable aggregate populations at 0.1 and 0.01 mM concentrations than R129385. Nonetheless, as evident from the graphs in

Figure 16, there was a concentration-dependent threshold around 0.01 mM below which all studied compounds formed aggregates with R_h values below 50 nm. RCpH analysis presented in Figure 16 corresponds to the most representative cases of NNRTI

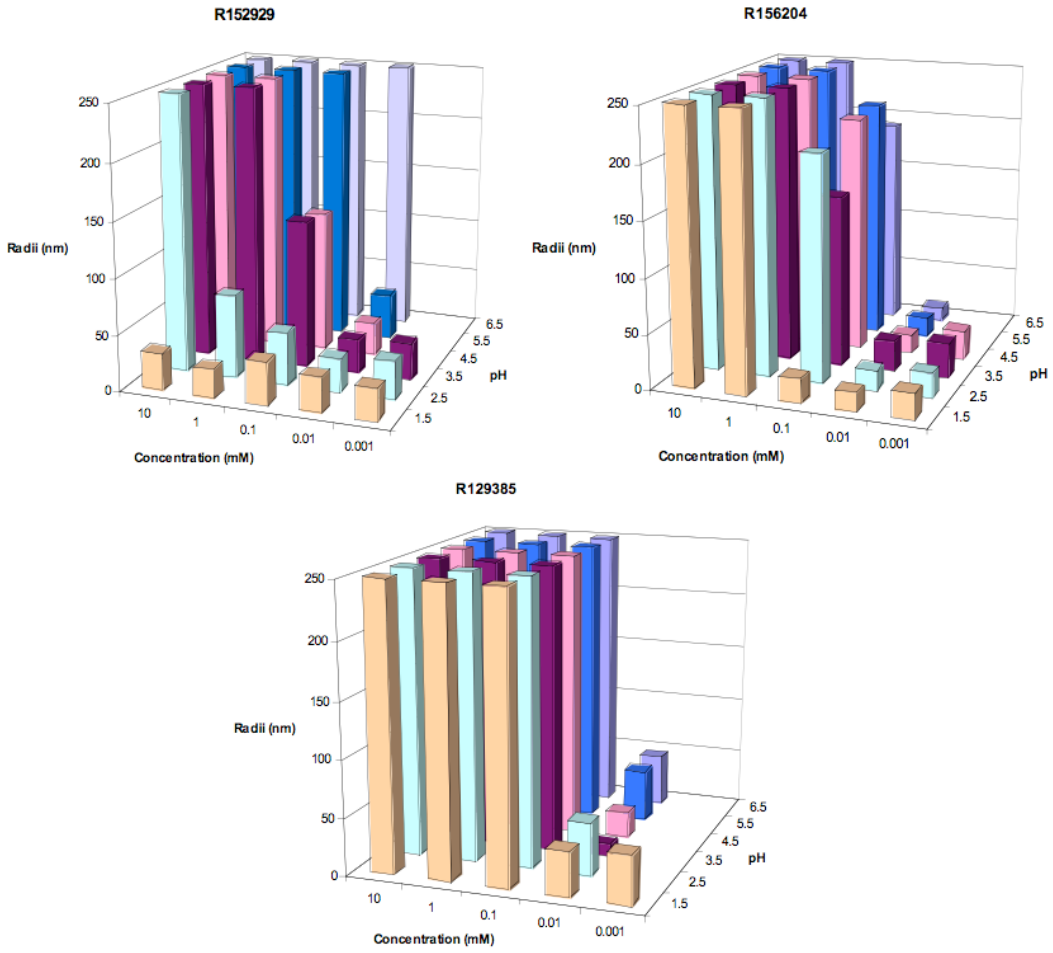


Figure 16. Concentration and pH dependence of NNRTI drug aggregate sizes (RCpH plots). RCpH plots illustrating the combinatorial effects of concentration and pH changes on aggregate size distribution for compounds with different bioavailabilities. Compounds R152929, R156204, and R129385 have good, intermediate, and poor bioavailability respectively. The readings below 0.001 mM were not reproducible due to instrument limitations.

behavior, yet the full subset of studied molecules is larger then shown (for similar analysis on other NNRTI molecules see Figure 17).

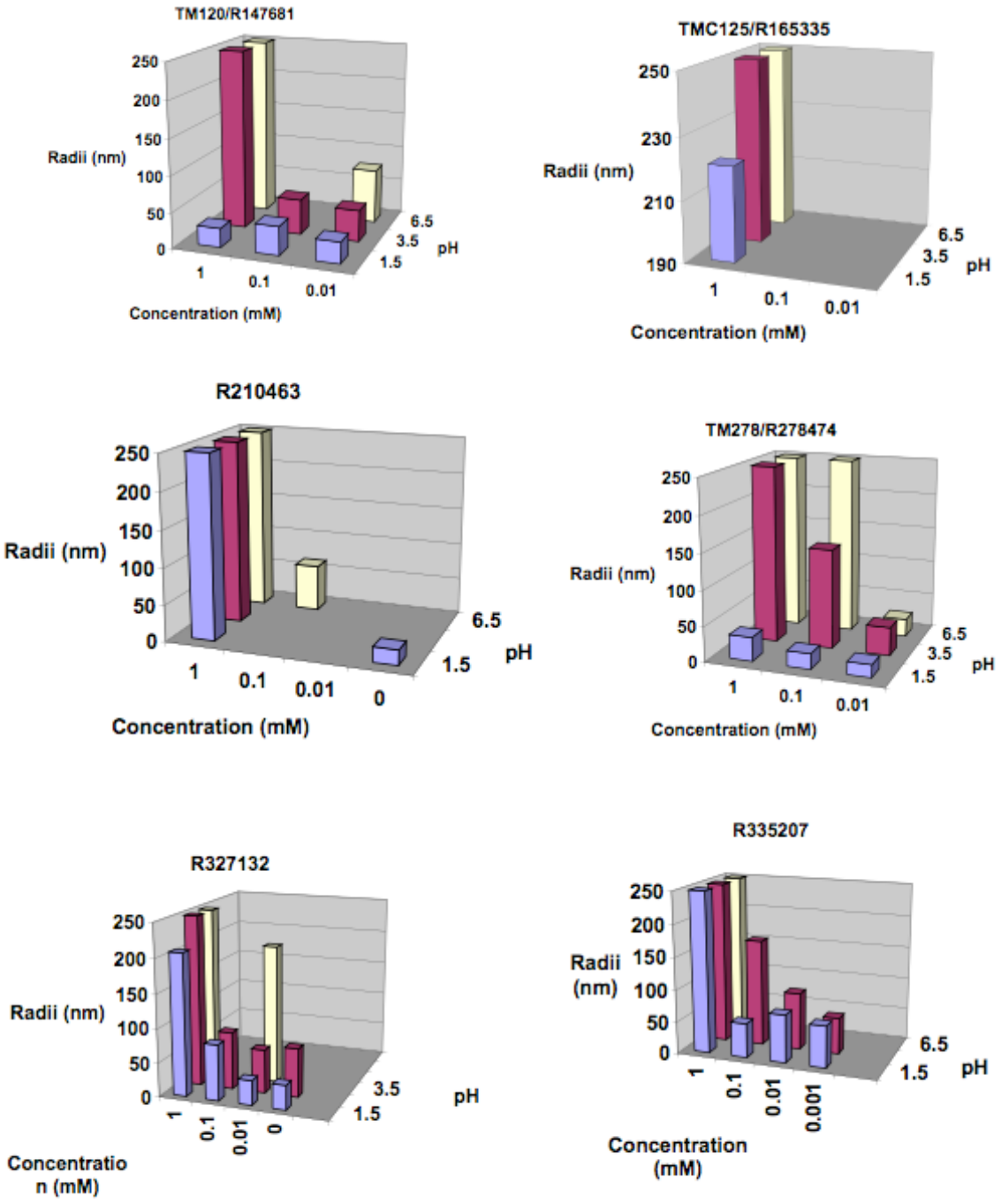


Figure 17-1. RCpH plots for 6 DAPY NNRTIs. Data collection was performed at pH 1.5, 3.5, 6.5 only. Each data point was collected in triplicate. Due to low scattering, reading at low drug concentration were sometimes not obtainable. Data collection was conducted with assistance from an undergraduate student, Tiffany Tsay.

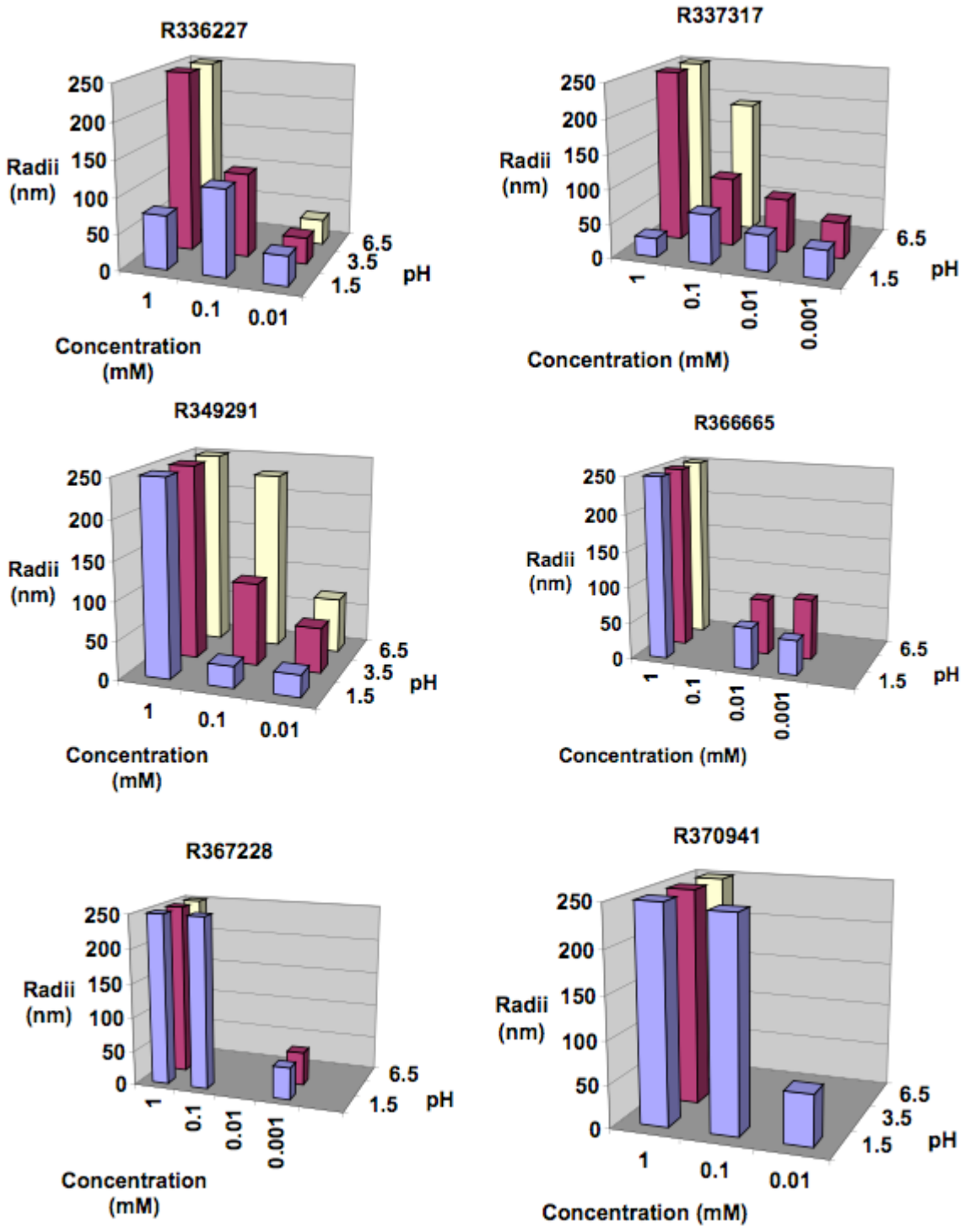


Figure 17-2. RCpH plots for 6 DAPY NNRTIs. Data collection was performed at pH 1.5, 3.5, 6.5 only. Each data point was collected in triplicate. Due to low scattering, reading at low drug concentration were sometimes not obtainable. Data collection was conducted with assistance from an undergraduate student, Tiffany Tsay.

Discussion

The physicochemical properties of hydrophobic compounds evaluated in this study displayed behavior consistent with Dr. Paul Janssen's hypothesis of NNRTI aggregate formation and also provide evidence of surfactant-like behavior of NNRTI molecules under acidic conditions in simple solutions. During the course of this study, a number of reports have appeared documenting that hydrophobic drug candidates can form aggregates in aqueous solutions (Feng et al., 2005; Feng and Shoichet, 2006; McGovern et al., 2002; McGovern et al., 2003; Pacheco and Carmona-Ribeiro, 2003; Seidler et al., 2003). McGovern *et al.* established that the aggregation behavior of hydrophobic drugs was responsible for promiscuous behavior in high-throughput screening for new lead compounds. It was proposed that hydrophobic compounds form aggregates capable of reversibly sequestering proteins in a non-specific manner leading to spurious identification of active inhibitors (hits) in high-throughput assays (McGovern et al., 2003). The authors developed a detergent-dependent screening tool that can distinguish whether a compound is a specific or non-specific protein inhibitor by evaluating inhibition of β-lactamase as a function of surfactant addition to experimental conditions. Those compounds that did not show inhibition and formation of drug aggregates in the presence of a detergent were identified as non-specific inhibitors that display promiscuous inhibitory activity only while in an aggregated form (Feng et al., 2005; Feng and Shoichet, 2006). In our work we independently developed DLS and EM analytical methods to evaluate hydrophobic aggregation behavior of NNRTIs in GI tract mimicking conditions and emphasize correlations between aggregation and drug delivery.

Aggregation of NNRTI compounds in simple solutions had been shown to be characteristic of non-promiscuous hydrophobic drug aggregators with amphiphilic properties. Simple solutions studies had shown that NNRTI aggregation is dependent on drug concentration, solution pH, and ionic strength while being relatively insensitive to surfactant. As described by Feng et al, hydrophobic drug aggregation can be linked to nonspecific inhibition of the drug-like molecules. However, based on a large subset of experimental studies (in vitro, in vivo, and structural), we know that NNRTI compounds are specific inhibitors of HIV-1 RT that form colloidal nanostructures under physiologically relevant conditions. Although the binding of NNRTI molecules in the present study is clearly specific to HIV-1 RT, we believe that detergent-sensitivity of aggregation behavior observed by Feng et al., can be used not only to determine specificity of binding but also to predict and evaluate the hydrophobic drug absorption properties of hydrophobic drug candidates. Formation of homogeneous aggregate population in the context of a detergent has been extensively used in hydrophobic drug formulation and delivery design. We hypothesized that formation of colloidal aggregates in the absence of surfactant is driven by the intrinsic properties of the drug molecules. Therefore we believe that drug molecules' ability to form small colloidal structures under conditions of relatively high drug concentration in a detergent-insensitive manor can be attributed to their surfactant-like characteristics.

Also some trends could be seen relating the compound pKa values and aggregation(Lewi et al., 2004): the three DAPY derivatives with the three highest pKa values (pKa>5; compounds TC120, R152929, and TMC278) formed aggregates and had very good rat and human absorption. The five compounds with the lowest rat AUC

values all had pKa values 3.5 or lower and, as stated above, four of the five formed larger structures in the pH experiments (Table 1 and Figure 16). Given the hints from EM that larger structures might form from the coalescence of smaller aggregates, these data may imply that the higher pKa values, which would correspond to more complete ionization (by protonation of pyrimidine nitrogens in this case) might permit stabilization of the aggregates with radii 30-100 nm under acidic conditions that may be important for efficient absorption after oral administration. These observations are confined to the DATA/DAPY series studied in this work and different classes of hydrophobic compounds are expected to have their own structure-aggregation correlations.

Absorption of NNRTI aggregates

Observed aggregation tendencies of NNRTI compounds in simple solutions offered support for the prediction that the presence of the NNRTI aggregates favors their lymphatic absorption via M cells in MALT of the GI tract. As has been established by Dressman *et al.*, poorly soluble compounds (such as in BCS Class II and IV) have dissolution rate-limiting absorption profiles that can be affected by the GI tract environment (Dressman and Reppas, 2000; Martinez and Amidon, 2002; Neuhoff et al., 2003). It has also been proposed that solubility plays a crucial role in hydrophobic drug absorption rate. A transfer model has been proposed suggesting that high drug concentrations can lead to precipitation during transit across the pH gradient of the GI tract that can be correlated with poor oral absorption (Kostewicz et al., 2002; Kostewicz et al., 2004). Lymphatic absorption of drug molecules is not a new concept, however, the mode of transport proposed in our model is. Most lymphatic drug absorption paradigms

described to date are based on ability of lipophilic drugs to solubilize in the chylomicron fraction of the lymph (Charman and Stella, 1986; Charman et al., 1992). Our model suggests a parallel mode of absorption which is specific to those lipophilic drugs, such as the NNRTIs studied here, that have the ability to form small aggregated structures independent of either natural triglycerides or formulation additives acting as lipid vehicles. Developed model integrates available information to explain the unique absorption properties of highly hydrophobic compounds.

For the NNRTIs used in our study, oral bioavailability in human subjects was established by administration of 100 mg doses and determination of corresponding absorption (AUC) values. Assuming an immediate release formulation in a fasting individual and stomach volume in the range of 50 to 250 ml, the approximate gastric concentration of NNRTIs with molecular masses between 300 and 500 Da can be estimated to vary roughly from 1 to 10 mM. At high drug concentration and pH, our data indicated a sharp transition from aggregates to larger structures, likely representing a phase transition to precipitate formation. It was also observed that at concentrations below 0.01 mM, compound aggregation became concentration and pH independent and the aggregate size for all compounds was consistently smaller than 100 nm in diameter (see Figure 48). In conjunction with the earlier findings by Desai *et al.*, that polymeric particles with diameter ~100 nm are absorbed more readily in Peyer's patches than larger particles (Desai et al., 1996), it may be that all NNRTIs at concentrations below 0.01 mM would form aggregates of appropriate size that can be absorbed via the M cell route. In addition to this observation it should be stressed that there were also some NNRTIs that were able to form aggregates at concentrations above 0.01 mM, suggesting that those

compounds could be absorbed efficiently by M cells at drug concentrations higher than 0.01 mM.

In addition to the supportive findings on the possible M cell absorption of hydrophobic inhibitor aggregates, we also observed correlations between the *in vitro* generated aggregation of NNRTIs and their oral bioavailability in rat and man. Under simulated gastric conditions of low pH (1.5) and drug concentrations ranging from 1 to 10 mM, inhibitors with low absorption formed large structures with radii greater than 250 nm, while compounds with better absorption formed smaller aggregates with smaller dimensions. The aggregate size and good absorption correlation did not hold true when aggregation was measured at the NNRTI concentrations below 0.01 mM because all NNRTIs formed aggregates with radii equal or less than 50 nm. Based on our analysis, *in vivo*/*in vitro* correlation properties for NNRTIs can be summarized by the following statement: "NNRTIs capable of forming aggregates with radii less than or equal to 50 nm at concentrations above 0.01 mM, at various physiologically relevant pH points, *in vitro*, are more likely to have good oral bioavailability."

Location-dependent hydrophobic drug absorption model

Based on the aggregate lymphatic absorption model and correlations established in this study, we proposed a GI location and aggregate size dependent model for absorption of hydrophobic drugs (Figure 18). The physiology of the GI tract is such that gastric absorption of many drugs is low due to inability to cross the mucus lining of the stomach (Forssell, 1988). A locus of high absorption in the GI tract is at the upper small intestinal region called the duodenum. Although the duodenum is a favorable absorption

site, the drug residence time in this region of the GI tract is comparably short, therefore, regions of absorption such as the colonic area, where the rate of absorption is low but the residence time is significantly longer, is of interest as well (Camilleri et al., 1989; Waterman and Sutton, 2003). Based on our physicochemical findings, highly hydrophobic drug candidates in immediate release formulations can form aggregates in the stomach environment and some drugs can form small aggregates even at relatively high concentrations. If aggregates of the appropriate size can be maintained with increasing pH during the transition through the GI tract, larger numbers of aggregates

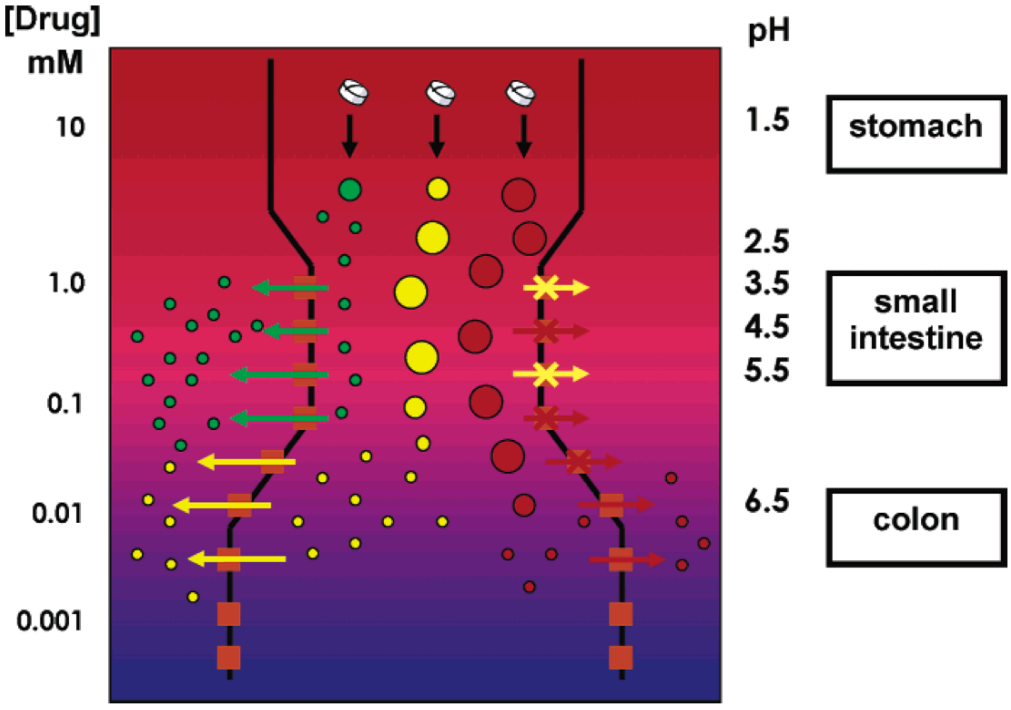


Figure 18. Location-dependent hydrophobic drug absorption model. Bold lines represent the GI tract: stomach, small intestine, and colon. Orange squares represent various absorptive cells in the GI tract including M cell and enterocytes. The color gradient reflects the pH changes in respective sections of the GI (red – low pH, blue – high pH). Rough estimates of pH and concentration are shown on the right and left sides of the image, respectively. Color assignment: red: poor absorption (AUC in rat/human <1 μg·h/mL); yellow: intermediate absorption (1 μg·h/mL <AUC in rat/human <5 μg·h/mL); green: good absorption (AUC rat/human >5 μg·h/mL).

potentially would enter the duodenum for absorption by M cells (and potentially other absorptive cells including enterocytes). Assuming that the aggregate lymphatic absorption model is correct, the presentation of a larger number of drug aggregates of appropriate size to the duodenal surface would result in high drug absorption and hence, high absorption. On the other hand, compounds that are unable to form large numbers of aggregates at high drug concentrations in the stomach would miss the duodenal absorption opportunity and only at ~0.01 mM concentration and below would be able to form aggregates with radii less than 100 nm as the predominant species. Formation of aggregates for these drugs would take place at the later stages of GI transit and would result in presentation to less efficient absorption sites in the system, which could explain the low absorption properties associated with these compounds.

Materials and Methods

DLS

Data collection was performed using the DynaPro-MS800 dynamic light scattering/molecular sizing instrument (Protein Solutions, 1994 - 2000) with argon laser wavelength λ=830 nm, a detector angle of 90°, and typical sample volume of 20 µl. Each light scattering experiment consisted of 20 or more independent readings, each 10 seconds in duration. Data analysis was conducted using DynaPro Instrument Control Software for Molecular Research DYNAMICS™ (Version 5.26.60). For very large particles with radii greater than 250 nm, full data collection was unfeasible, therefore,

less than 20 measurements were acquired and size determination was based on empirical data readings. Each DLS experiment was repeated in triplicate. To minimize dust interference in light scattering experiments, all stock solutions were filtered with 0.22 μm MillexGV sterile filters.

All NNRTIs were dissolved in either 100% DMSO or 100% PEG400 and stock solutions at 50 or 20 mM concentrations were kept frozen at -80°C. Inhibitors were diluted from DMSO stocks into solutions mimicking GI conditions of a fasting individual, vortexed for at least 10 seconds, and then evaluated by DLS. Gastric conditions *in vitro* were simulated by using: a) 0.15 M HCl, pH 1.5; b) 0.15 M HCl, pH=1.5 with 0.1% (v/v) tyloxapol (surfactant); and c) 20 mM malonic acid, pH 1.5. Note that although the 0.15 M HCl, pH 1.5 condition showed formation of NNRTI aggregates, it was a highly unstable system for DLS evaluation, and therefore was replaced with 20 mM malonic acid buffer at pH 1.5 which produced aggregates of similar size, gave more stable DLS readings, and permitted buffering over a broad range of relevant pH ranges. The solutions intended to mimic pH transitions in the GI tract were 20 mM malonic acid for pH ranging from 1.5 to 4.5 and 50 mM malonic acid for pH ranging from 5.5 to 6.5.

Aggregate thermal stability was evaluated for five NNRTIs at 4°C, 25°C, and 37°C at 0.4, 0.7, and 1.0 mM drug concentrations. Sample preparation and data collection was done in the similar fashion to that described earlier.

Malonate buffers were prepared using malonic acid and sodium malonate solutions (J.T.Baker, Inc.). Tyloxapol (Sigma-Aldrich), a nonionic surfactant from the Triton group, was used to evaluate detergent-dependent aggregation behavior of TMC278. Tyloxapol micelles with characteristic 3.5 nm hydrodynamic radii (Regev and

Zana, 1999) were observed by DLS in control solutions and were clearly distinguished in experimental conditions from drug containing aggregate populations. TMC278 was kindly provided by Janssen Pharmaceutica.

TEM

Inhibitors were diluted from 20 or 50 mM stocks into 0.15 M HCl, pH 1.5, with and without 0.1% tyloxapol to various concentrations: 0.1 mM, 0.25 mM, 0.5 mM, and 1 mM. At room temperature, 5 μl samples of the reaction mixtures were applied to the carbon-coated grids and negatively stained with 2% uranyl acetate. Micrographs were taken on a JEOL 1200 EX transmission electron microscope at 80 kV.

Chapter 2. Conformational variability and amphiphilicity of DAPY NNRTIs in small molecule crystal structures

Synopsis

Here is presented the small molecule structure analysis for five DAPY NNRTIs with differing abilities to inhibit a broad range of HIV-1 RT mutants and oral bioavailability. Analysis of the small molecule crystal packing established that some of the DAPY NNRTI molecules are amphiphilic, and therefore sample both hydrophobic and hydrophilic types of interactions. This finding is important for our understanding of aggregation properties of the drug molecule suggesting that some of the NNRTIs have better ability to sample surface-active properties than others. Structure analysis also provided unequivocal evidence of multiple conformational sampling by DAPY NNRTIs via torsional variability that had been previously hypothesized to play major role in mechanisms by which NNRTIs can overcome resistance mutations. Furthermore, small molecule structures revealed sampling of an additional order of freedom associated with interconversion of hybridization states at linker nitrogens in the selected drug molecules. Sampling of alternative hybridization states by linker nitrogen atoms was determined to be a mechanism for preservation of hydrogen bond networks in the crystals. Similar hydrogen bond motifs were observed in RT/NNRTI complexes suggesting that the drug molecules might be sampling alternative hybridization states in the bound form as well. Also, weak polar interactions between the cyano group of the cyanovinyl moiety of the TMC278 and the aromatic rings of the symmetry-related molecules was detected. Similar interactions had been observed in the RT bound TMC278 structures. The weak,

non-specific nature of this interaction revealed the breadth of conformational freedom available to TMC278 in the hydrophobic tunnel of HIV-1 RT.

Introduction

Aggregation properties of NNRTI compounds reported in chapter one suggest that NNRTI aggregate formation depends on the atomic composition of the drug molecules. In the series of 15 DATA/DAPY NNRTIs studied, the five molecules that contain oxygen atoms (either as aryl ring linkers or as hydroxyl group) had the lowest exposures in rat (see Figure 4, Table 1) and four of the five formed larger structures (R_h >250 nm) in the pH 1.5 condition in the absence of tyloxapol (Figure 7) condition, as opposed to the other ten that in each case formed aggregate structures with radii 30-100 nm. These oxygen atoms are able to act as hydrogen-bond acceptors even at pH 1.5, allowing for different contact patterns than nitrogen or carbon atoms.

Although the precise mechanism of aggregate formation is unknown to us, we expect that the process is driven, among other not yet identified factors, by burial of hydrophobic surface area (as with the hydrophobic effect in the formation of oil droplets in water or the folding of proteins in aqueous environments), which may balance the energy cost incurred by forming cavities in the water. The arrangement of molecules within the aggregates may resemble an environment anywhere along the continuum from liquid-like to liquid crystals to solid-like in which microcrystalline domains may be present. The ways in which the hydrophobic molecules will pack in the aggregates will depend on solution conditions (pH, buffer, salt, ionic strength, etc.) and the detailed structures and properties of the molecules, including the number of polar atoms, number of hydrogen bond donors and acceptors, conformational flexibility, and presence of ionized groups, among other factors. Depending on the solvent, salts, and other additives

(e.g., tyloxapol or PEG400) in solution, additional molecules and species may be present within the aggregates.

To address the structure activity relationship for DAPY NNRTIs, we studied the crystal structures of five NNRTI compounds with differing binding and oral bioavailability properties. DAPY NNRTI molecules have characteristic structural features of two hydrophobic aryl rings (wings I and II) and a pyrimidine ring (P) (see example in Figure 19). It has been earlier hypothesized and experimentally supported, in the structural studies of DAPY NNRTIs in complexes with HIV-1 RT, that they achieve multiple conformations by altering dihedral angles τ1 and τ2 responsible for 'wing I' and 'wing II' rotation relative to each other (Das et al., 2008; Das et al., 2004). The flexible

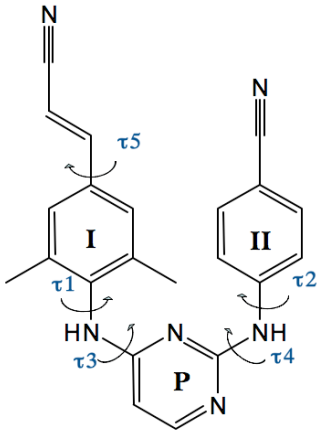


Figure 19. NNRTI nomenclature. As an example of a DAPY NNRTI, the structure of TMC278 is presented with wing I, wing II and pyrimidine rings labeled as I, II, and P respectively. Black arrows correspond to rotations associated with torsion angles τ1 – τ5.

nature of the molecule was proposed to play an essential role in improving ability of the NNRTI molecules to inhibit a broad range of RT mutants. Rotation of torsional angles τ1 and τ2 was predicted to involve relatively low energy barriers and therefore were

suggested to afford broad conformational sampling (Das et al., 2004). Dihedral angles τ3 and τ4 were observed to be less dynamic while τ5 was shown to explore varying conformations selection of which was mostly influenced by interaction with critical residues in the pocket of the protein comprising the hydrophobic tunnel.

Results

Here are reported the small molecule X-ray structures of TMC278 in protonated, bis-DMSO, and anhydrous forms (Figure 20); and TMC125 in an anhydrous form (Figure 21). In our analysis and discussion we also include a set of three DATA/DAPY NNRTI small molecule structures: R120393, R152929, and R185545, kindly provided by The Center for Molecular Design, Belgium (Figure 22). All TMC278 crystal forms have one molecule per asymmetric unit, though the bis-DMSO crystal has two alternative conformations, A and B, with partial occupancy of 0.72 and 0.28 for atoms C(12) through C(22); and 0.909(2), 0.091(2) for DMSO molecules A and B respectively (Figure 20). Data collection and refinement are reported in Table 2, and complete data on bond distances and angles is included in Figure 23.

Conformational and torsion angle variability analysis

Out of five available small molecule structures of DATA/DAPY NNRTIs, four displayed a “horseshoe” or a “butterfly” conformation where |τ3| and |τ4| are less than 90° (Figure 21). One exception to this conformational trend was R120393, with |τ3| at 178° resulting in a drug molecule conformation transformed into a twisted or an extended

one, otherwise known as a “seahorse” (Figure 22). Small molecule crystal structure conformations of drug molecules are similar to those previously observed in the corresponding NNRTI complex structures with HIV-1 RT (Das et al., 2008; Das et al., 2004). Although the overall horseshoe shape of the drug molecules in the small molecule X-ray structures is mostly conserved, the intrinsic flexibility of the molecules can be appreciated from the number of different conformations determined for one of the NNRTI molecules, TMC278 (Figure 23). Torsional variability is analyzed by comparison of torsion angle sampling for all NNRTI structures (Table 3). For torsion angles $\tau 1$ and $\tau 2$ the standard deviation is 35° and 43°, respectively, while for $\tau 3$ and $\tau 4$ these values are 15° and 14° for horseshoe conformations. If the seahorse conformation was also included in the calculation the standard deviation for these angles increased to 64° and 49°, respectively. Predominance of one general fold suggested that the “horseshoe” conformation has lower energy than the “seahorse”.

Torsional flexibility in the small molecule crystal structures of DAPY NNRTIs can be also evaluated in terms of sampling of alternative hybridization states at the linker nitrogens, N(3) and N(4) of the drug molecules. We analyzed planarity at the linker nitrogens in all available small molecule crystal forms (see Methods) and our observations suggest both linker regions are capable of sp2 and sp3 hybridizations (Table 4). The change in hybridization is associated with inter-conversion between planar and tetrahedral conformations. Table 4 summarizes hybridization information for three TMC278 crystal forms as well as TMC125, R120393, R152929, and R185545. TMC125 and R185545 were not analyzed using earlier described method due to the absence of hydrogen atom information. As shown in Table 4, the linker I nitrogen, N(4), observes

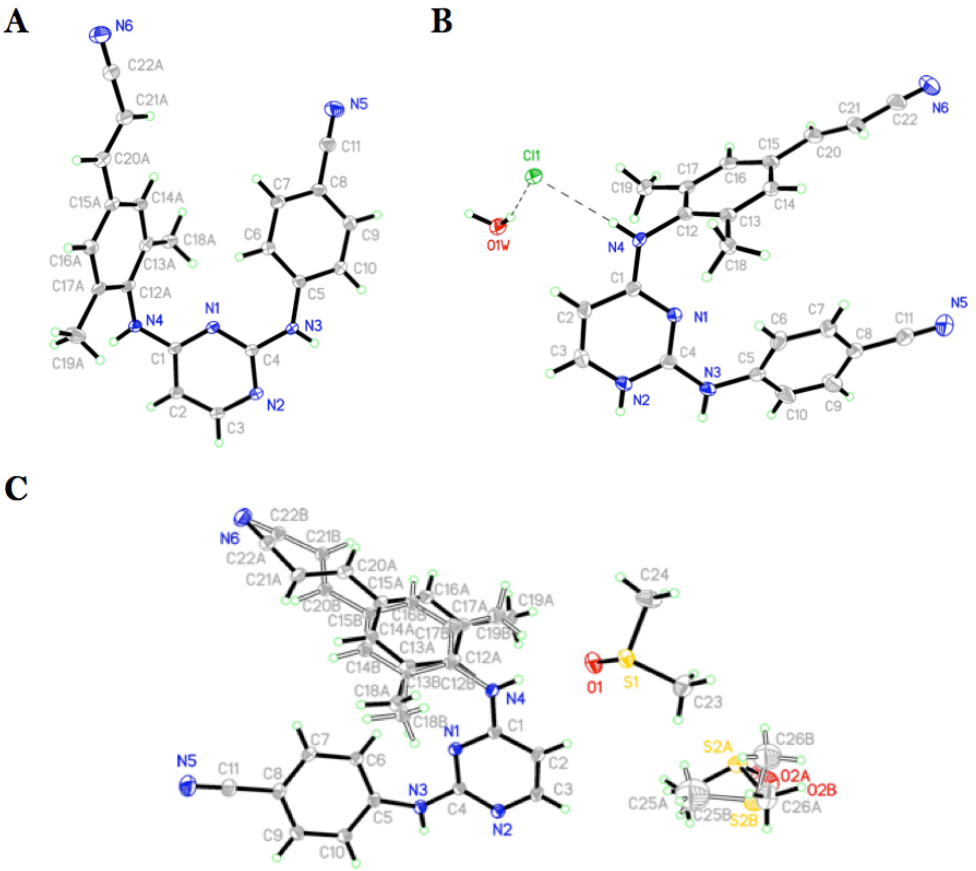


Figure 20. Thermal ellipsoid representations of TMC278 small molecule X-ray structures. Thermal ellipsoid plots of TMC278 crystal structures in anhydrous (A); bis-DMSO (B); and protonated (C) forms.

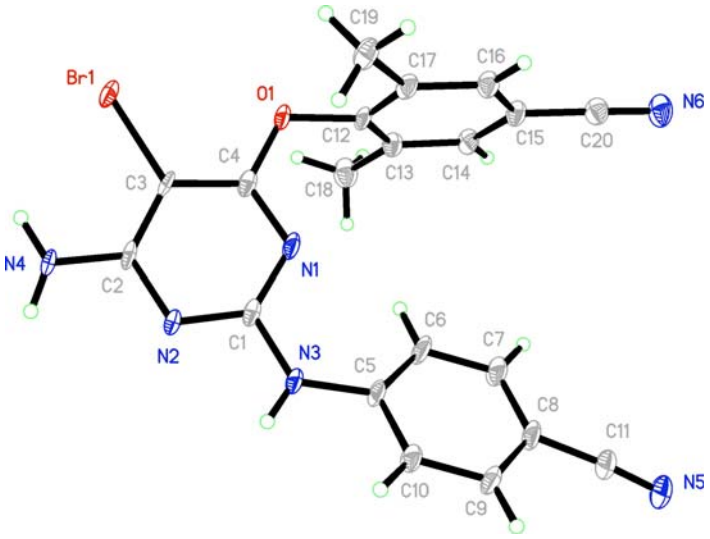
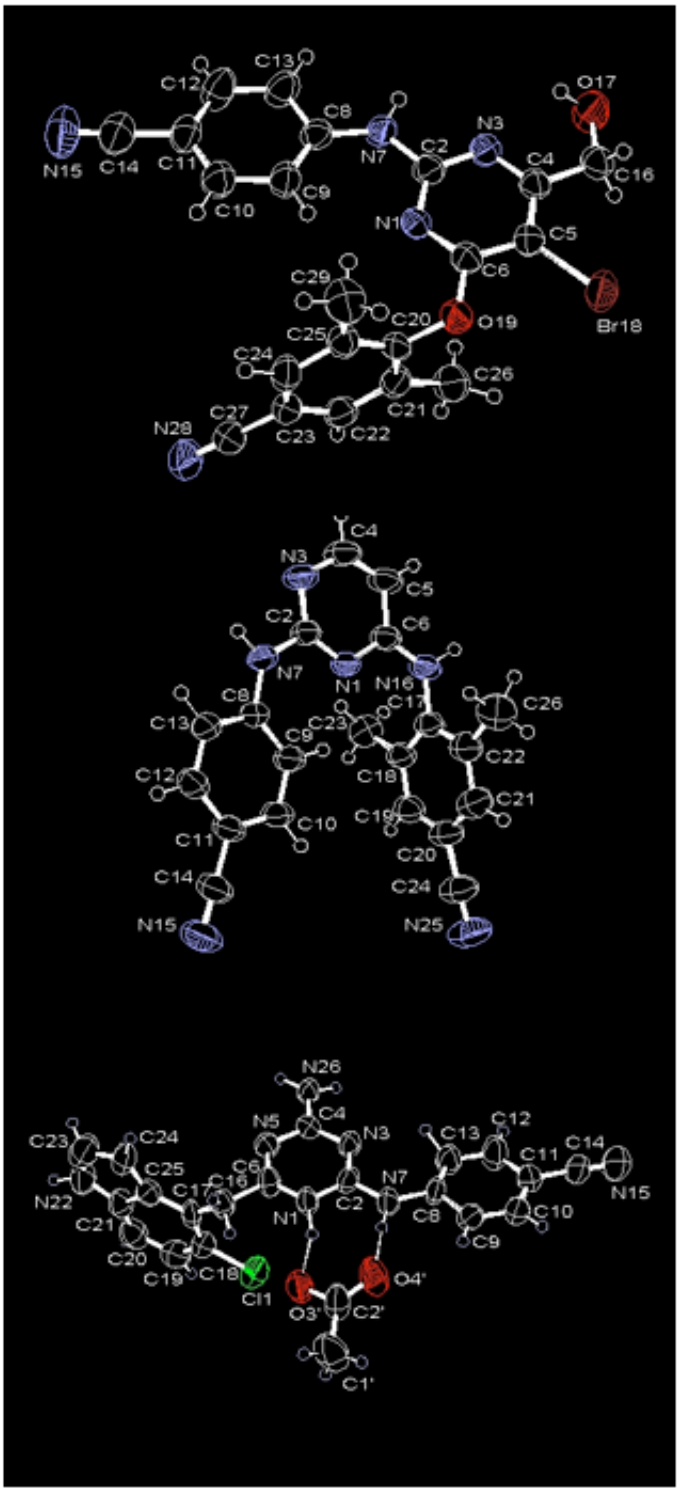


Figure 21. Thermal ellipsoid representation of TMC125 X-ray structure.



R185545

R152929

R120393

Figure 22. Thermal ellipsoid representation of R185545, R152929, and R120393 structures. Coordinates and thermal ellipsoid plots were provided by the Center for Molecular Design, Belgium.

	Anhydrous	bis-DMSO	Protonated	TMC125
Crystal data				
Composition	C22H18N6	C26H30N6O2S2	C22H21ClN6O	C20H15BrN6O
MW (g/mol)	366.42	522.68	420.9	435.29
Temperature (K)	100(2)	100(2)	100(2)	100(2)
Crystal System	Monoclinic	Triclinic	Triclinic	Triclinic
Space group	$P2_1/c$	$P\bar{1}$	$P\bar{1}$	$P\bar{1}$
Radiation source	Mo Kα	Mo Kα	Cu Kα	Mo Kα
a (Å)	8.3666(7)	8.4047(7)	8.8181(3)	7.3486(12)
b (Å)	13.8607(11)	11.9357(10)	9.6228(4)	8.3803(14)
c (Å)	15.6024(13)	13.6480(11)	13.7870(5)	16.974(3)
α	90	79.390(2)	98.111(2)	77.413(3)
β	91.765(2)	89.015(2)	97.811(2)	84.212(3)
γ	90	80.080(2)	111.599(2)	69.815(3)
V (Å^3)	1808.5(3)	1325.44(19)	1054.207(7)	957.2(3)
Z	4	2	2	2
Dx (Mg/m^3)	1.346	1.31	1.326	1.51
m (mm^{-1})	0.084	0.236	1.816	2.171
Shape, color	paddle, colorless	sword, colorless	plate, colorless	plate, colorless
Dimension (mm^3)	0.51 x 0.16 x 0.05	0.50 x 0.17 x 0.04	0.14 x 0.06 x 0.025	0.36 x 0.11 x 0.02
Θ range (°)	1.97 – 30.46	2.10 – 30.53	3.31 – 58.89	2.46 – 28.28
Data Collection				
h	-11→11	-11→11	-9→9	-9→9
k	-19→18	-16→16	-10→10	-10→11
l	-22→22	-19→19	-15→15	-22→22
Total no. of reflections	20727	15698	8925	10011
No. of independent reflections	5429	7945	2920	4707
R_{int}	0.0387	0.0283	0.027	0.0362
% completeness at θ_{max}	98.9	98.1	96.5	99.3
Refinement				
R[I>2σ(I)]	0.0528	0.0545	0.0384	0.0484
R_w	0.1269	0.133	0.0909	0.1175
S	1	1.008	1.001	1.009
No. of parameters	325	393	301	264
No. of restraints	0	49	0	0
$\Delta\rho_{max}$ (eÅ^{-3})	0.509	0.509	0.261	2.31
$\Delta\rho_{min}$ (eÅ^{-3})	-0.182	-0.182	-0.222	-0.945

Table 2. Crystallographic data collection and refinement statistics for single crystal X-ray structures. Listed information is for anhydrous, bis-DMSO, and protonated forms of TMC278, and TMC125 in the anhydrous form.

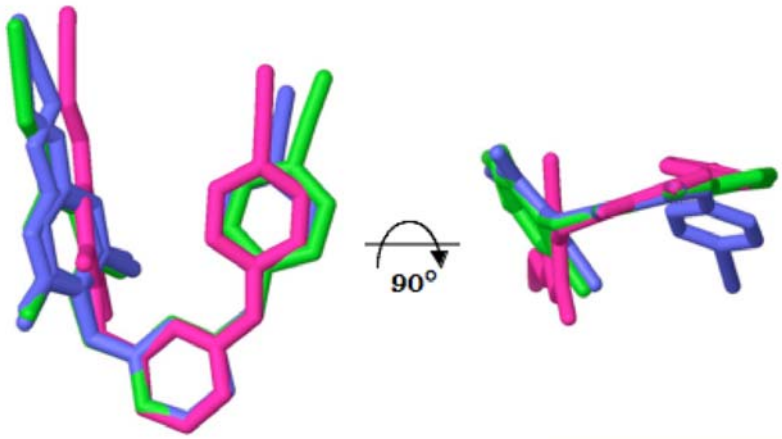


Figure 23. Superimposed view of anhydrous, bis-DMSO, and protonated X-ray structural forms of TMC278. Superposition was done using non-hydrogen atoms of the pyrimidine rings. Color assignment: green: anhydrous; blue: bis-DMSO; pink: protonated.

		τ1	τ2	τ3	τ4	τ5
TMC278	protonated	31	-81	2	13	-11
	bisDMSO-A	76	35	-18	-22	-11
	bisDMSO-B	66	35	-9	-22	174
	anhydrous	106	-10	12	6	14
TMC125	anhydrous	97	30	2	5	n/a
R120393	acetate	9	62	179	130	n/a
R152929	anhydrous	103	8	4	-8	n/a
R185545	anhydrous	86	3	28	2	n/a

Table 3. Dihedral angle measurement comparison in small molecule X-ray structures of DAPY NNRTIs. Angles τ1 – τ4 were identified using the following convention: τ1 – C(1)-N(4)-C(12)-C(13); τ2 – C(4)-N(3)-C(5)-C(6); τ3 – N(1)-C(1)-N(4)-C(12); τ4 – N(1)-C(4)-N(3)-C(5). Angle τ5 (C(14)-C(15)-C(20)-C(21)) was measured only for TMC278; see Figure 19.

sp3 hybridizations at a higher frequency than the linker II atom, N(3). That being said, the only NNRTI structures where linker I atoms displayed planar orientation were TMC125 and R185545 where instead of nitrogen this position was occupied by oxygen. Comparison of linker II orientation in different DAPY NNRTIs was more exhaustive in this study that that of linker II because atomic composition of the later varied. As reported in Table 4, the sp2 hybridization was observed in two out of five linker II atoms in NNRTI structures (protonated TMC278 and anhydrous TMC125). In the example of the protonated TMC278 structure, it was

	TMC278				TMC125	R120393	R152929	R185545
	Protonated	bis-DMSO A	bis-DMSO B	Anhydrous				
Stacking Rings*	0	0	0	0	1	3	0	4
Hydrogen Bonds	3	3	3	2	4	4	2	0
Linker I Atom	N	N	N	N	O	C	N	O
General Shape	U	U	U	U	U	E	U	U
Hybridization	sp3	sp3	sp3	sp3	sp2**	sp3**	sp3	sp2**
H4N out of plane distance (std) [Å]	-0.2935 (0.0244)	-0.3503 (0.0213)	-0.2385 (0.0215	0.3466 (0.0187)	n/a	n/a	-0.2251 (0.0154)	n/a
Out of Plane Angle [°]	21.4	26.4	17.4	24.4	n/a	n/a	15.0	n/a
Linker II Atom	N	N		N	N	N	N	N
Hybridization	sp2	sp3		sp3	sp2	sp2**	sp3	sp2**
H3N out of plane distance (std) [Å]	-0.0734 (0.0250)	0.3505 (0.0200)		-0.1425 (0.0180)	0.0193 (0.0388)	n/a	-0.1133 (0.0149)	n/a
Out of Plane angle [°]	4.8	26.8		9.4	1.3	n/a	7.3	n/a

Note: greater than 5 sigma difference is considered significant
* - stacking interactions were accounted for if the plane-plane distance was 3.5 ± 0.5 Å
**- carbon and oxygen hybridizations were not effected

Table 4. Hybridization switches in small molecule X-ray structures of DAPY NNRTIs. The linker region geometry comparison was conducted for TMC278 in protonated, bis-DMSO, and anhydrous forms, and for the TMC125, R120393, R152929, and R185545 structures.

possible to determine the reasons behind alternative hybridization sampling (Figure 24A). The HN4 atom deviation from the C(12)-N(4)-C(4) plane is most likely driven by

hydrogen bond formation that coordinates chloride ion and water by the symmetry-related drug molecules (see Figure 24B). The extent of deviation can be interpreted using the displacement distance from planarity and the angle of rotation necessary for planarity.

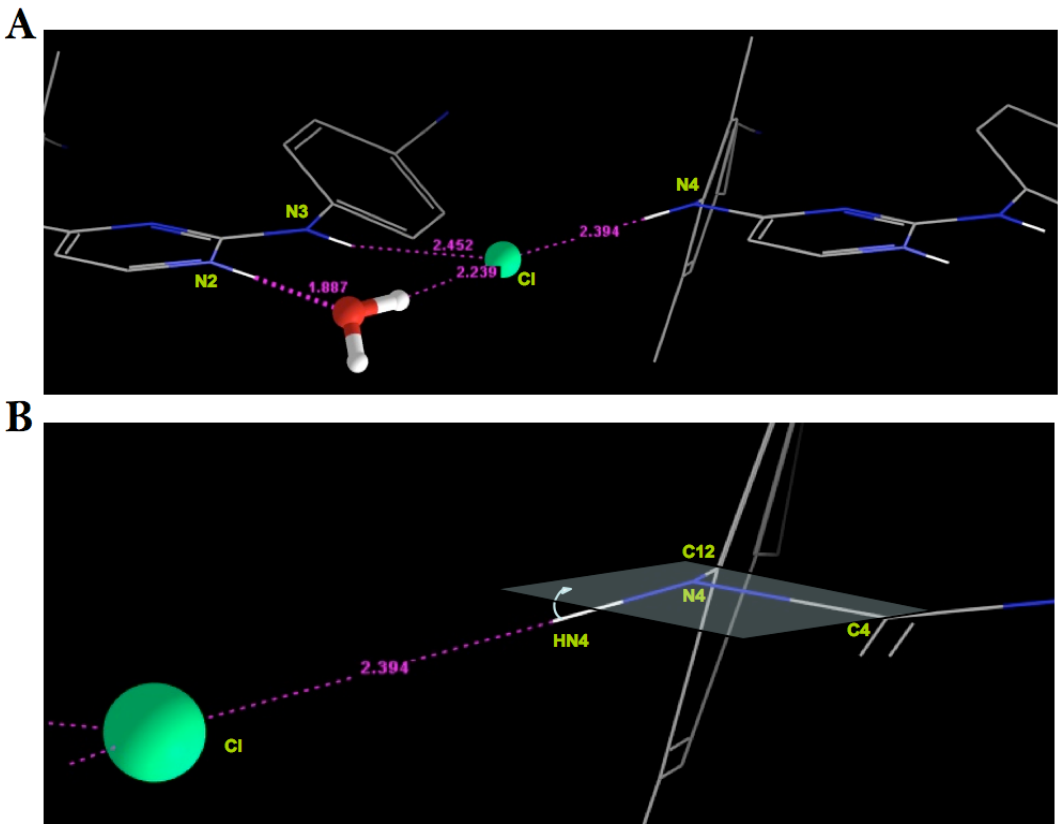


Figure 24. Sampling of alternative hybridization states at the linker I and II nitrogens in protonated TMC278 structure. A. Hydrogen bond network between two symmetry related protonated TMC278 molecules with ordered solvent. B. Close-up of hydrogen bonding between linker I hydrogen, HN4 and chloride ion. Light blue arrow reflects the necessary rotation for HN4 to be in the plane C12-N4-C4. Color assignment: green: chloride; red: oxygen; blue: nitrogen; white: hydrogen; gray: carbon. Note: only polar hydrogens are displayed. Hydrogen bonds are shown in magenta with corresponding distances in Å.

In the case of the protonated TMC278 structure, displacement from the C(12)-N(4)-C(4) plane for linker I hydrogen is 0.3 Å while the deviation angle is at 21°. On the other hand, linker II hydrogen (HN3) has only a minor deviation from the plane described by a

4.8° rotation and a 0.07 Å displacement. Due to the fact that both parameters are less than five standard deviations, linker II hydrogen is considered to be planar and therefore sp2 hybridized.

Hydrogen bonding analysis

Hydrogen bond networks were also evaluated for DAPY NNRTI small molecule structures. Analysis involved evaluation of crystal packing, bond distances, and bond angles. Conserved preference for hydrogen bond formation were observed between linker HN(3) and HN(4) atoms and symmetry-related molecules and/or ordered solvent, either chloride ions as in the protonated structure; or a symmetry-related pyrimidine ring N(2) atom and an ordered DMSO oxygen, in bis-DMSO structure; or pyrimidine N(2) and cyano N(5) atoms in the anhydrous form (Table 5). Among all hydrogen bonding networks observed in the TMC278 structures, these interactions had the most optimal geometries (bond angles closest to 180° and hydrogen bond length of 2.5 ± 0.25 Å) and therefore were considered to be significant (Figure 25, Table 5). In the case of TMC125 neither linker nitrogen N(3) nor oxygen at the linker I position formed hydrogen bonds with surrounding molecules. Besides the linker regions, the pyrimidine ring moiety of the molecules was also noted to form hydrogen bonds with either symmetry-related drug molecules or ordered solvent. These interactions were strong with good geometry (Table 5), however generally weaker than those formed by linker nitrogens, with one exception in TMC125 where the linker nitrogens were not involved in hydrogen bond formation and the best hydrogen bonding geometry was that of pyrimidine nitrogen N2.

	d(H…A) [Å]	D(D...A) [Å]	<(DHA) [°]
Anhydrous			
N(3)-H(3N)...N(2)#1	2.47(2)	3.337(2)	172(2)
N(4)-H(4N)...N(5)#2	2.26(2)	3.089(2)	163(2)
bis-DMSO			
N(3)-H(3N)...N(2)#1	2.23(2)	3.0289(19)	173(2)
N(4)-H(4N)...O(1)	2.14(2)	2.9341(18)	168(2)
Protonated			
N(3)-H(3N)...Cl(1)#2	2.45(3)	3.294(2)	161(2)
N(4)-H(4N)...Cl(1)	2.40(3)	3.228(2)	178(2)
N(2)-H(2N)...O(1W)#1	1.88(3)	2.721(3)	177(2)
TMC125			
N(4)-H(4NA)...N(2)#1	2.24(4)	3.088(4)	166(3)

Table 5. Selected intermolecular hydrogen bond distance and angle information from anhydrous, bis-DMSO, and protonated TMC278 small molecule crystal structures. #1 and #2 identify symmetry-related molecules. Symmetry operations performed to generate symmetry-related molecules and the full list of hydrogen bonds identified in small molecule structures are listed in Table 8-3.

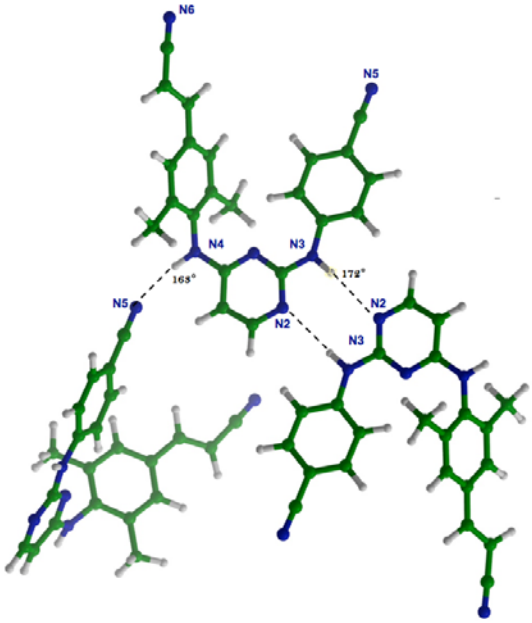


Figure 25. Hydrogen bonding conserved network motif in TMC278 X-ray structures. Sample packing corresponds to the anhydrous form of TMC278. Color assignment: green: carbon, blue: nitrogen, grey: hydrogen.

Among available DAPY NNRTI structures the most extensive information on drug/solvent interactions could be obtained from the protonated structure of TMC278. The ordered solvent in the protonated structure of TMC278 formed a solvent channel that clearly showed formation of a hydrogen bond network with pyrimidine rings and linkers of drug molecules while their hydrophobic moieties are buried away. Figure 26 depicts

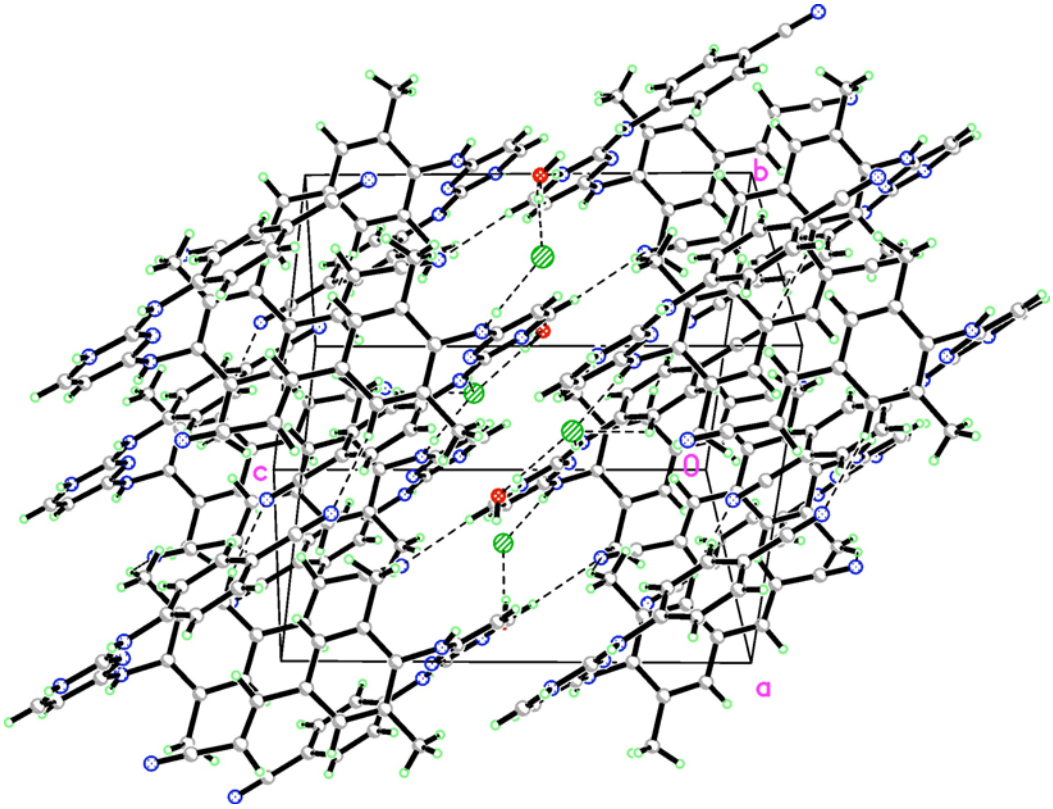


Figure 26. Protonated TMC278 small molecule crystal packing. Color assignment: white: carbon; blue: nitrogen; solid green: chloride; solid red: oxygen; clear green: hydrogen.

the crystal packing of the protonated TMC278. Based on the drug and ordered solvent hydrogen bonding network it can be concluded that TMC278 molecules have amphiphilic tendencies. Burial of the wings and extensive hydrogen bonding interactions of the pyrimidine adds directionality and distinction between the head and the tail of the

molecule. The fact that similar modes of hydrogen bonding pattern were observed in both, protonated and neutral structures, suggested that amphiphilicity was not limited only to the protonated TMC278 molecules, yet the presence of the charge probably amplified that property.

To evaluate the effect of protonation on the structure of DAPY NNRTI molecules we compared the bond distances of anhydrous, bis-DMSO, and protonated structures of TMC278. Figure 27 shows root square deviation (RSD) of bond distances in bis DMSO and protonated TMC278 structures versus the same distances in the anhydrous TMC278 structure. The deviation distribution highlights the largest variation in bond distances in

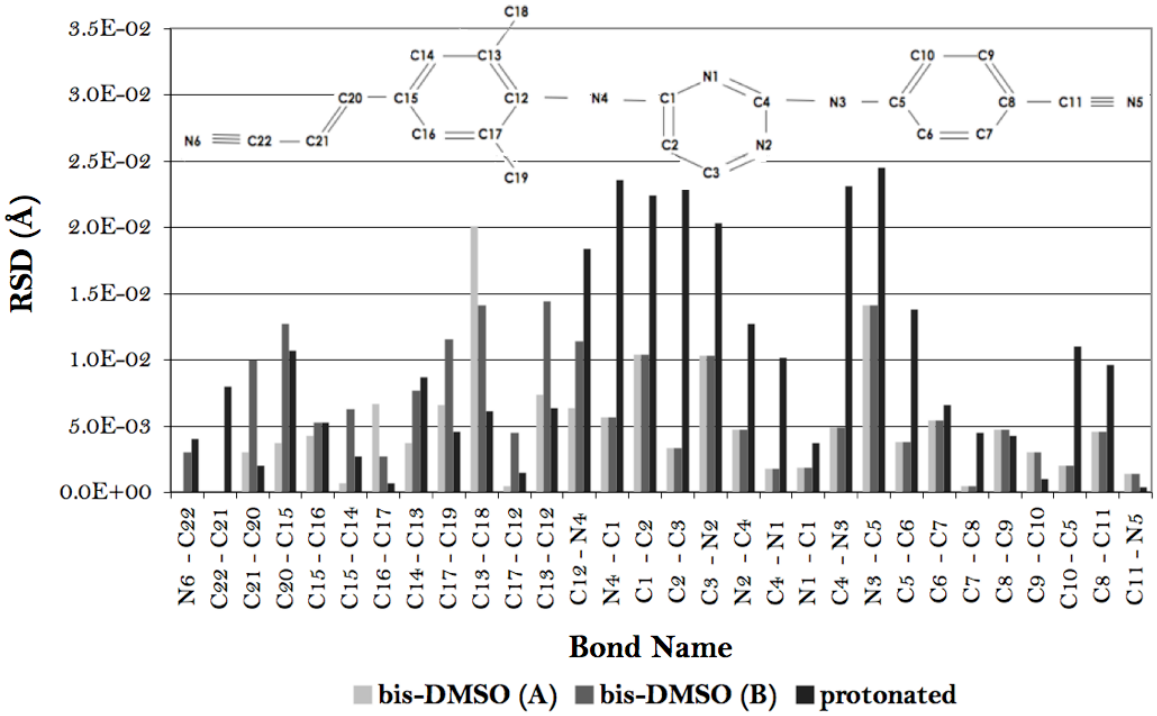


Figure 27. Plot of RSD for non-hydrogen atom pair distances in the three small molecule X-ray structures of TMC278. The RSDs were calculated between bis-DMSO or protonated forms and anhydrous TMC278 structures. Color assignment: light grey: bis-DMSO (A); dark grey: bis-DMSO (B); black: protonated TMC278.

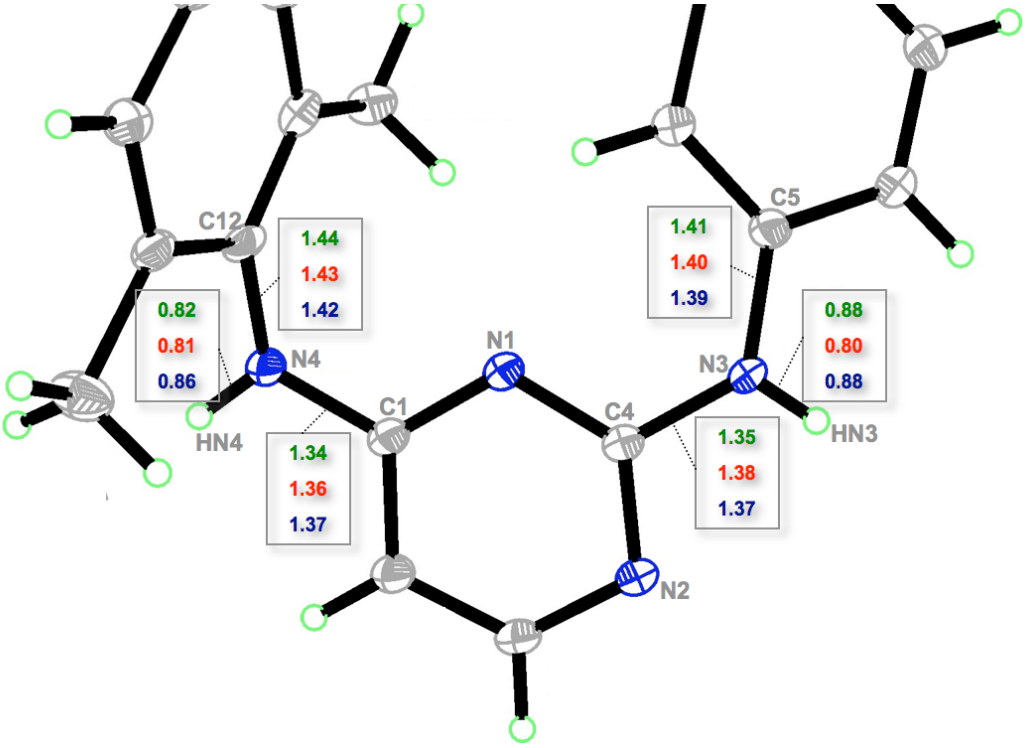


Figure 28. Comparison of bond distances (in Å) for protonated, bis-DMSO A, and anhydrous TMC278 small molecule structures in the linker region. Color assignment: green: protonated; red: bis-DMSO; blue: anhydrous.

the three structures. The most substantial differences between these structures lie in the pyrimidine ring [C(1)-C(2), C(2)-C(3), C(3)-N(2)] and linker regions [C(12)-N(4), N(4)-C(1), C(4)-N(3), N(3)-C(5)] of the protonated TMC278, whereas bis-DMSO A/B had comparably small deviations from corresponding anhydrous TMC278 non-hydrogen atom bond distances. That being said, almost all bonds with exception of C(21)-C(22), C(11)-N(5), and C(22)-N(6), showed variation above one standard deviation (>0.003 Å). Tighter hydrogen bonding to chloride ions in the protonated structure translated into a substantial change of the bond distances for pyrimidine ring bonds, linkers and wing I of the molecule, while the cyano groups and wing II were not affected by these long range interactions. Though the cyano group in the anhydrous TMC278 structure formed

hydrogen bonds, the distance of the triple bond was not significantly affected. Distance comparison of the linker-regions' associated bond distances for C(1)-N(4); N(4)-C(12) and C(4)-N(3); N(3)-C(5) depicted in Figure 28, suggested that protonation and the presence of ordered chloride ion, resulted in shortening of C(1)-N(4) and C(4)-N(3) by 0.03 and 0.02 Å respectively, simultaneously being compensated by elongation of C(12)-N(4) and C(5)-N(3) by 0.2 Å. As discussed before, the linker nitrogens were involved in coordination of chloride ion and therefore the differences in the bond lengths of atoms associated with the linker can be explained in terms of these interactions. Interestingly, the protonation of N(2) affects the bond distances on the pyrimidine in an uneven fashion, in which the N(2)-C(3) bond was more significantly affected than the bond length of N(2)-C(4). This uneven distribution might be related to the change in electron composition and distribution in the structure.

Crystal packing analysis

In addition to analyzing the conformational flexibility and hydrogen bond motifs, we also evaluated hydrophobic interactions in the crystal structures of DAPY NNRTIs. Figure 29 shows the crystal packing samples from five DAPY NNRTI small molecule X-ray structures. Evaluation of crystal packing was based on four parameters: ring-to-ring stacking distance (d_s), angle between a stacking rings pair (<s), buried surface of stacking between two stacking rings expressed as solvent excluding surface area of stacking ($SESA_s$), and the total number of ring-to-ring stacking interaction accounted for individual drug molecules (Table 6, details in Methods). Optimal parallel displaced stacking interaction between two rings was defined by d_s of 3.5 ± 0.5 Å, <s of 180 ± 20°

of 0 ± 20°, and $SESA_s$ >20 $Å^2$ (see Methods for details). In order to compare stacking interactions between different crystal structures, a quality of stacking (QS) parameter was introduced to reflect on the number of descriptors of parallel stacking interaction falling in the optimal range for a given stacked pair of rings.

Among five studied compounds, some had extensive stacking interactions while others failed to form any. Among analyzed DAPY NNRTI small molecule structures, only two compounds had all three rings involved in parallel displaced stacking resulting in π–π interactions, R120393 and R185545 (Table 6). Even though both of these compounds employed all three rings in ring-to-ring stacking interactions, the strength or quality of stacking for these molecules was fundamentally different. Though stacking distances in R120393 packing were within the optimal range of 3.5 ± 0.5 Å, the two stacking angles $<s_{(II...P)}$ in R120393 were determined to be suboptimal (<180°) producing only 45 $Å^2$ of total buried surface area between the stacked rings. In contrast to R120393, R185545 had shown better geometry in both ring-to-ring distances and angles, therefore generating a total of 66 $Å^2$ of solvent excluded surface attributed to ring-to-ring stacking (Table 6). On the other hand, there were also structures that had not formed optimal aromatic ring stacking interactions: protonated and bis-DMSO TMC278 forms, and R152929. Potential stacking interactions were observed in the bis-DMSO TMC278 structure between symmetry-related wing I and pyrimidine ring, however the distance of separation and $SESA_s$ suggested that orientation was not optimal for π–π stacking interactions. The TMC278 stacking interactions were critically affected by the presence of ordered solvent and protonation of the drug molecules. In the case of the anhydrous TMC278 structure, one clear stacking interaction between wing II and a symmetry-

related wing II ring was observed. Based on the ring-to-ring distance of 3.8 Å, co-planar orientation (characterized by a 180° angle), and a buried surface of 27 $Å^2$ the stacking was determined to have optimal geometry. In contrast to the anhydrous TMC278

	Ring Pair	d_s (Å)	<s (°)	$SESA_s$ ($Å^2$)	QS	N_{TS}
TMC278/Anhydrous	II…II	3.8	180	27	**3**	
	II…P	5.2	170	0	1	**1**
	I…I	5.5	180	0	1	
				Total=27		
TMC278/bis-DMSO	I…I	4.5	180	10	1	0
	II…P	4.4	142	14	0	
				Total=24		
TMC278/Protonated	I…I	4.8	180	12	1	0
	II…P	5.5	170	0	1	
				Total=12		
R120393	I…II	3.9	170	22	3	3
	(II…P)×2	3.8	160	23	2	
				Total=45		
R152929	I…I	4.4	160	21	1	0
	I…II	5	180	0	1	
	II…P	4.6	165	11	1	
				Total=32		
R185545	I…I	4	180	21	3	3
	II…II	3.6	168	25	3	
	II…P	3.8	170	20	3	
				Total=66		
TMC125	I…I	3.7	180	25	3	1
	II…P	4.2	160	13	0	
	II…II	4.2	180	20	2	
				Total=58		

Table 6. Crystal packing analysis for five DAPY NNRTI small molecule structures. I, II, and P correspond to wing I, wing II, and pyrimidine ring respectively. Ds – stacking distance, <s –angle between two stacking rings, $SASE_s$ – buried surface of stacking between selected rings, QS – quality of stacking interactions (1 point for each descriptor), N_{TS} – total number of stacking interactions for an individual drug molecule.

structure, bis-DMSO and protonated drug forms did not form optimal stacking

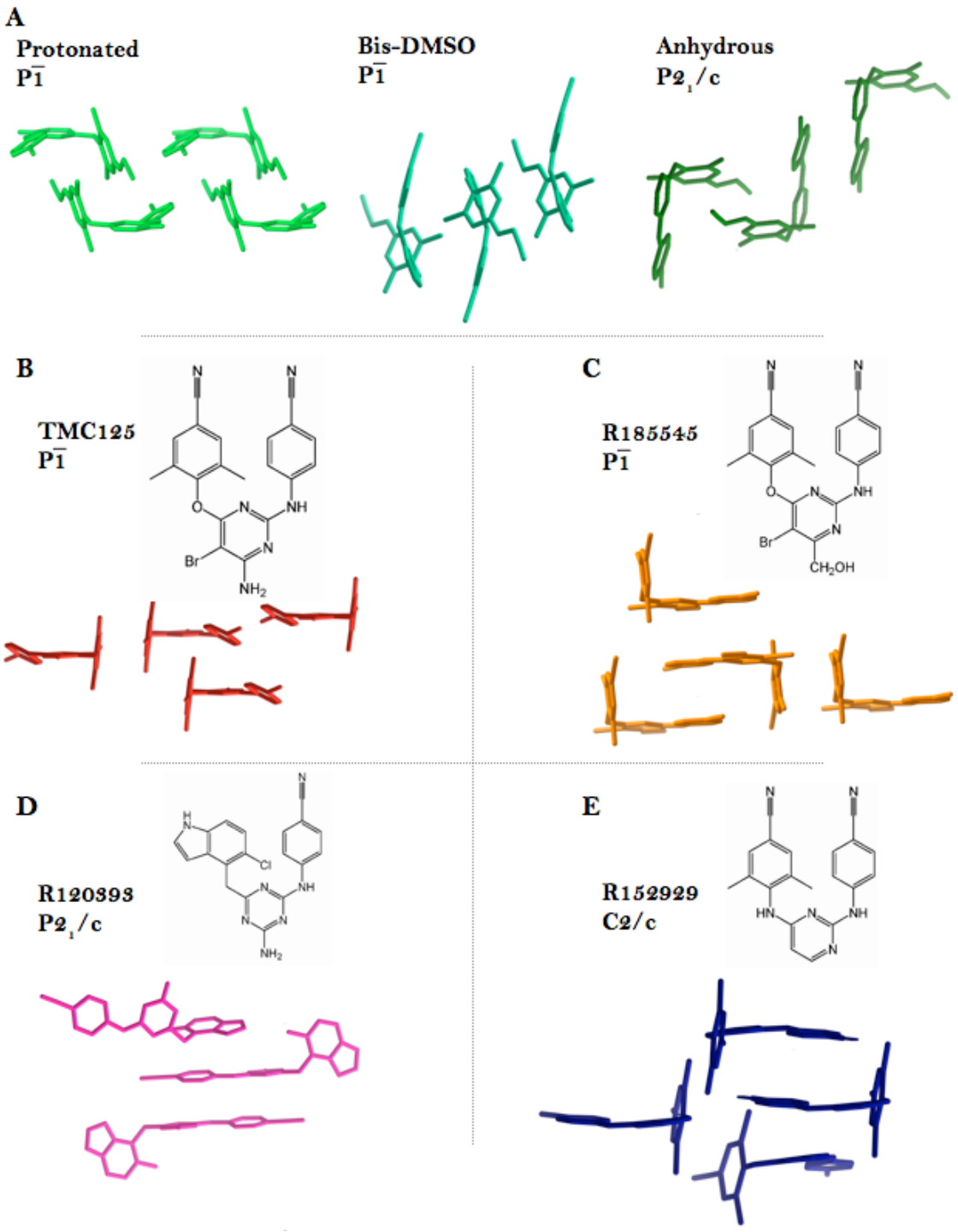


Figure 29. Packing analysis of five DAPY NNRTI molecules. **A.** TMC278 **B.** TMC125 **C.** R185545 **D.** R120393 **E.** R152929. Symmetry-related molecules were generated using symmetry operators appropriate for the corresponding space group.

interactions between any of the possible ring pairs. As shown in Figure 30, the total buried surface area for an individual molecule in the context of the crystal packing was determined to be the largest for the anhydrous crystal form (223 $Å^2$), in comparison to bis-DMSO and protonated TMC278 structures where SESA values were 166 and 79 $Å^2$, respectively. These results suggest that the crystal packing of the anhydrous TMC278 form is tighter, and therefore produces more contact area between symmetry-related molecules while as in bis-DMSO and protonated crystal forms the packing is increasingly less compact. In the anhydrous structure the most interactive part of the molecule is that of wing II, whereas in bis-DMSO more contacts occur with the wing I moiety.

Though different in many ways, all three TMC278 crystal forms pack their cyano groups against symmetry-related molecules in a similar fashion. From the hydrogen bond analysis discussed earlier, it was established that the wing II cyano group formed hydrogen bonds with pyrimidine rings of symmetry related TMC278 molecules only in

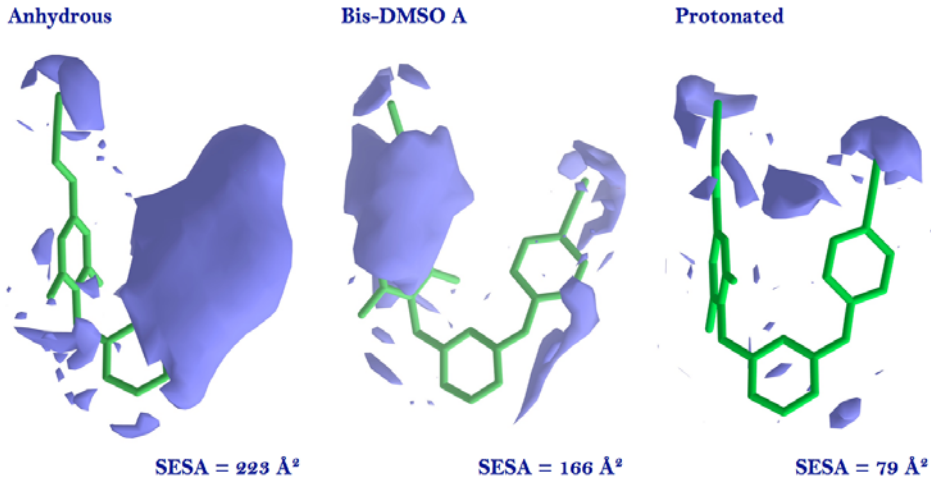


Figure 30. Buried surface analysis for TMC278 small molecule X-ray structures. Crystal structures are organized in order of decreasing total buried surface expressed as $SESA_s$.

the anhydrous structure while as in the others TMC278 crystal forms such clear hydrogen bond formation was not observed. On the other hand, the cyanovinyl moiety of the TMC278 molecule was observed to display conserved through-space electronic interactions between the cyano group and aromatic rings of symmetry-related molecules in all three structures. Figure 31 shows cyanovinyl moiety packing modes in three TMC278 crystal forms. The distance between cyanovinyl groups and ring centers varied from 3.0 Å in the protonated, 4.1 Å in the anhydrous, to 5.5 Å in the bis-DMSO structure while orientation of the cyanovinyl moieties relative to the plane of the symmetry-related wing I in all cases was coplanar. In the anhydrous and the protonated crystal forms, the cyano group of the cyanovinyl moiety made close contacts with pyrimidine rings of the symmetry-related molecules. In the bis-DMSO structure, the two wing I phenyl rings formed a parallel-displaced stacking interaction, forcing the cyanovinyl moiety to make exclusive packing contacts with the pyrimidine ring of a symmetry-related molecule (Figure 31). The distance between cyano groups of the cyanovinyl and pyrimidine rings were optimal for electronic exchange at 3.1 Å, 4.0 Å, and 3.8 Å in protonated, bis-DMSO, and anhydrous crystal forms, respectively. The extent of contact made between cyanovinyl moiety and aromatic rings was evaluated using the buried surface comparison. Cyanovinyl-to-aromatic rings buried surface ranged from 13 to 19 $Å^2$ (Table 7). Out of three analyzed structures in two (protonated and anhydrous), cyanovinyl group made contacts with two aromatic rings: the phenyl ring of wing I and the pyrimidine ring (P). As discussed earlier, due to the wing I – wing I parallel-displaced stacking alignment the buried surface between cyanovinyl moiety and aromatic rings in this structure was smaller than in the other structures. In all three structures, the

pair-wise contact surface area comparison had shown that there was more surface interaction between cyanovinyl groups and P ring than with the phenyl ring of wing I.

The buried surface results for small molecule X-ray structures were compared to those observed in X-ray structures of TMC278 in complex with wild and mutant HIV-1 RT (Figure 32). As depicted in the figure, the most extensive contacts were made by the cyanovinyl group in the hydrophobic tunnel of the binding pocket and aromatic residues Trp229 and Tyr188 (distances range from 3.9 to 4.3 Å). The nature of the packing was similar to that observed in the small molecule structure, with tangential orientation of the cyano group relative to the aromatic rings. Just as geometry, the buried surface area was

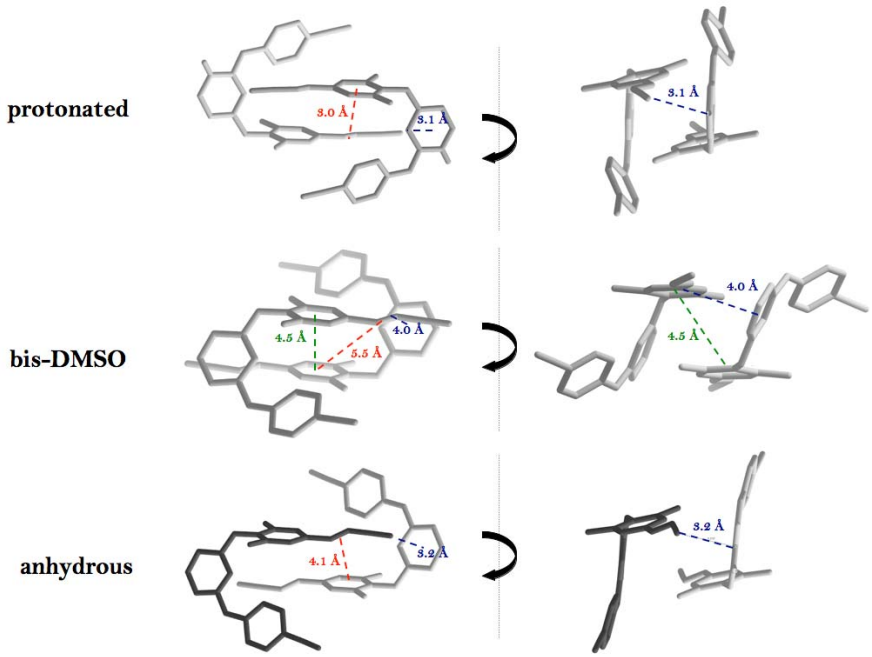


Figure 31. Cyanovinyl group stacking modes in TMC278 X-ray structures. Packing interactions are shown in two alternative views for protonated, bis-DMSO, and anhydrous TMC278 structures. Color assignment for mapped distances: green: ring center to ring center distance; red: ring center to the center of the vinyl bond; blue: pyrimidine ring center to cyano group.

TMC278 crystal form	SESA $_{\text{Wing I - Wing I}}$ (Å^2)	SESA $_{\text{CNV - Aromatic Ring}}$ (Å^2)
Protonated	53	R_I - 14 P - 16
Bis-DMSO (A)	42	P - 19
Anhydrous	58	R_I - 13 P - 16

Table 7. Buried surface analysis for TMC278 X-ray structures. Calculation of buried surface between two rings ($SESA_{WingI\text{-}WingI}$) was based on Wing I, cyanovinyl, and dimethyl phenyl (R_I) or pyrimidine ring (P) atoms; and buried surface between cyanovinyl and nearest aromatic ring ($SESA_{CNV\text{-}Aromatic\ Ring}$) included cyanovinyl and aromatic ring atoms were included.

also comparable between the small molecule structure and RT complexes. Between RT complex structures, the packing of the cyanovinyl moiety against hydrophobic residues was generally conserved yet the contact area reflecting the strength of the interaction differed. In the wild-type and in the L100I/K103N mutant structures, the contact area

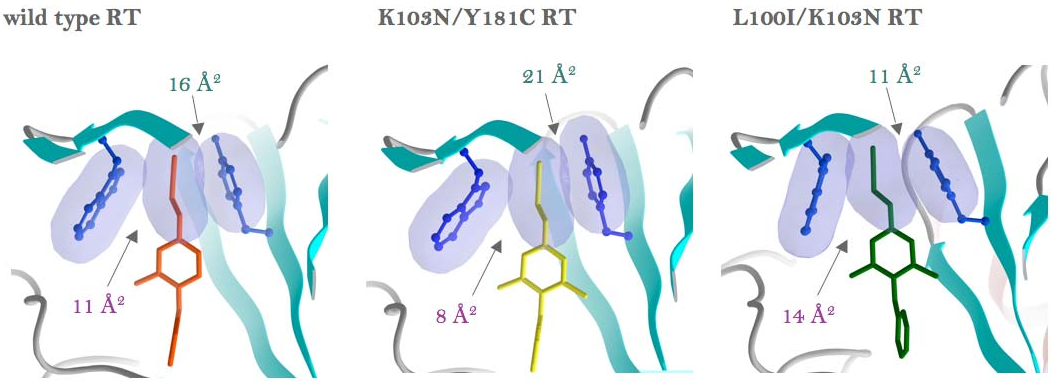


Figure 32. Buried surface area comparison for crystal structures of TMC278 in complex with HIV-1 RTs. The packing interactions between cyanovinyl moieties of TMC278 and hydrophobic residues in the wild-type and mutant RT complexes were estimated using buried surface calculations (see Methods).

between the cyanovinyl moiety and Tyr188 was greater than in the K103N/Y181C structure (16 $Å^2$ vs. 14 $Å^2$ vs. 21 $Å^2$) whereas the extent of packing between Trp229 and the drug was in reverse order. These observations suggested that although in general the packing of the cyanovinyl to aromatic rings was favored, increased availability of different aromatic residues in this location allows for the drug molecule to sample a larger number of comparable conformations.

Discussion

Some of the DAPY NNRTI molecules exhibit unique characteristics in both inhibition of their primary target and oral bioavailability. We have studied the small molecule X-ray structures of five DAPY NNRTI molecules with different binding and bioavailability characteristics and have looked for structural features that attribute to their exceptional abilities.

Drug aggregation

Exceptional oral bioavailability of some of the DAPY NNRTI has been linked to their corresponding aggregation properties in simple solutions (Frenkel et al., 2005). Aggregate formation mechanisms can be explored by a number of different methods, including small molecule X-ray structures. In the first chapter, *in vitro* and *in vivo* studies on DAPY/DATA NNRTI behavior in simple solutions had identified correlation between aggregation properties of these compounds and their oral bioavailability. We have proposed that molecules that are able to form large populations of aggregates with

~60 nm in diameter at relatively high drug concentrations and low solution pH are taken up by systemic circulation at high rate and therefore are orally more bioavailable.

The dependence on pH of the drug aggregate formation of the aqueous media suggests that the aggregate formation and stability are heavily dependent on the ionized drug population. The small molecule crystal structures of DAPY NNRTIs presented in this work provide evidence that some of the drug molecules in the protonated form display amphiphilic properties. In the TMC278 protonated crystal structure, most of the solvent exposed regions of the molecule correspond to pyrimidine rings, linkers, and cyano groups. Packing of molecules in the crystal and hydrogen bond networks with ordered solvent suggest that the molecule preferentially exposes the pyrimidine ring moiety while burying hydrophobic aromatic rings (wings I and II). If drug aggregate formation is considered in the context of micelle structure, surface molecules would be anticipated to be surface-active agents, i.e. surfactants. In drawing parallels between surfactants and TMC278 in the protonated form, the hydrophobic wings of the molecule behave analogously to the hydrophobic tail of a surfactant while the more polar pyrimidine ring and linker nitrogens can act as polar head groups. We hypothesize that aggregate formation is driven by hydrophobic interactions between drug molecules while water/aggregate interface formation is dependent on the amphiphilic, protonated species.

Of the five available structures, two molecules are in their protonated forms, TMC278 and R120393. As can be seen from the structures, the protonation sites for two molecules are different though both are localized at the pyrimidine ring atoms, N(2) and N(1) (Figures X2, X3). Between the two, the N(2) protonation site is only slightly less sterically hindered in TMC278 than in R120393, yet, the presence of the amino group at

C(2) position most likely disfavors protonation at the nearby nitrogens, therefore forcing protonation at the N(1) atom in R120393. The overall conformations of the two molecules are different; in the case of TMC278 it is a folded conformation (where $\tau 3$ and $\tau 4$ are less than 90°) while in the case of R120393 it is an extended conformation (where $\tau 3$ and $\tau 4$ are greater than 90°). Interestingly, both of the molecules have good absorption profiles in rats. Of the two, R120393 is more soluble in water, yet the TMC278 absorption profile is still exceptional. One possible explanation lies in the different paths of absorption taken by these analogous compounds. We hypothesize that R120393 absorption is enhanced by its increased, single molecule solubility while as in the case of TMC278 it is attributed to the structural feature that allows for more pronounced amphiphilic characteristics, such as separation between hydrophilic and hydrophobic moieties and amplification of surfactant-like features with preferential display of the amphiphilic molecules at the surface of the aggregate.

Biophysical data and computational studies (described later) suggest that the neutral molecules might occupy the internal structure of the drug aggregate. Small molecule crystal structures for molecules with different aggregation properties allows for comparison of the small molecule crystal packing that provides information on the extent of interaction between drug molecules. Of the five compared drugs, TMC278 in DMSO and R152929 have the least extensive hydrophobic interactions whereas R185545 has the most extensive network of π–π stacking rings. These results correlate with aggregate formation observations made in solution studies where R152929 and TMC278 has some of the best aggregate formation properties while as R185545 has the worst.

It is not known whether the drug aggregates have microcrystalline domains or fluid-like interior of an aggregate. Yet, the small molecule structure tells us that some of the drug molecules in their neutral form have higher propensity towards forming ring-to-ring stacking interactions while others do not. It is possible that compounds that are unable to form drug aggregates under any of the physiologically relevant concentrations have higher predominance of hydrophobic interactions than other analogs.

Protein binding

As inhibitors of HIV-1 RT, some of the DAPY NNRTIs are known to be exceptional in their ability to inhibit a broad range of known HIV-1 RT mutants. As proposed by Das, Lewi, *et al.*, the mechanism of overcoming NNRTI resistance mutations hinges on the ability of the drug molecules to attain multiple binding modes in the binding pocket (Das et al., 2004; Das et al., 2005; Lewi et al., 2003). The structural information obtained from the small molecule X-ray structures presented in this work supports this hypothesis and shows that multiple conformational modes of the molecules can be achieved not only via torsional changes at $\tau 1$ and $\tau 2$ dihedral angles responsible for rotations of the aromatic rings but also in the geometry of the linker atoms of the molecules. The structural changes observed for the drug molecules were clearly influenced by their corresponding environments, whether those were the binding pocket of the HIV-1 RT or the small molecule structures with different solvent content.

The small molecule structures reported here have shown that those DAPY NNRTIs molecules that have nitrogen atoms at the linker position acquire alternative geometries to accommodate conserved hydrogen bond motif formation. Even under

conditions of unfavorable geometry, hydrogen bonding was achieved by changes in hybridization of linker atoms (example of protonated TMC278). Nitrogen atom hybridization interconversion has been observed before in other systems and was shown to result in enhanced ability of the molecule to participate in hydrogen bond formation. For example, based on the crystallographic data analysis and *ab inito* molecular orbital calculations by Allen et al., intramolecular steric interaction was determined to facilitate conversion of nitrogen from sp2 to sp3 to promote hydrogen bond acceptor characteristic of the nitrogen in catalytic reaction (Allen et al., 1995). From the crystallographic studies conducted by Das *et al.*, we know that the N(3) linker atom forms hydrogen bonds with the carbonyl oxygen of K101, while N(4) participates in a hydrogen bonding interaction with carbonyl oxygen of E138 mediated through an ordered water molecule in the wild-type RT pocket (Das et al., 2008; Das et al., 2004; Das et al., 2005). Drawing parallels between the hydrogen bonding network in the binding pocket and information acquired from small molecule structures, we hypothesize that drug molecules with nitrogen linkers sample alternative hybridization states in the protein bound complex as well.

A number of studies have reported that changes in nitrogen atom hybridization can significantly affect drug molecule structure and drug/protein binding interactions. In some cases switch in hybridization was associated with increase in ligand entropy and decrease in affinity, meanwhile in others it facilitated enhanced binding selectivity through better accommodation of the target geometry (Allen et al., 1995; Brameld et al., 2008). Based on our observations, structural changes resulting from sp2 to sp3 conversion produce conformations with an additional order of freedom that are not observed in the landscape of motion for sp2 only atom-linker containing DAPY NNRTI

analogs. Knowing that some of the most potent compounds have nitrogen linkers, it is speculated that the additional order of freedom does not interfere with binding of the ligand. Moreover, hydrogen bond networks observed in the X-ray structures and those in the RT binding pocket show that presence of oxygen or nitrogen at these positions would allow for more stabilizing interactions in the protein pocket. Nitrogen is known to have stronger hydrogen-bond acceptor properties under the same conditions than oxygen (Bohm et al., 1996) and therefore would be anticipated to perform better in accommodating available hydrogen bond interaction in the everchanging RT binding pocket. Therefore in the mechanism by which DAPY NNRTIs overcome resistance mutations, it is hypothesized that ability of the molecules to preserve hydrogen bond interactions despite the changes the protein pocket landscape increases their potency. The hydrogen bond network is believe to act as an anchoring mechanism for the drug molecule while as the interconversion at the linkers allows to accommodate the newly acquired changes to the landscape. This observation suggest that hydrogen bond interactions via the linker atoms are fundamentally important to the ability of the drug to maintain strong affinity to various protein mutants.

Another insight into NNRTI/RT complex provided by the small molecule structures is a conserved packing motif of the cyanovinyl moiety with aromatic rings. As shown in Figure 31, we observed the cyanovinyl moiety of the TMC278 molecule forming conserved interactions with surrounding aromatic groups. As discussed in the review by Meyer *et al*, non-conventional interactions between polar groups and aromatic rings in small molecule structures are widely observed in X-ray structures and often are suggestive of small molecule binding properties in a protein target (Meyer et al., 2003).

Interactions between aromatic ring centers and nitrile or cyano groups identified in other molecular systems were referred to as direct HOMO-LUMO orbital interaction (Ramasubbu et al., 1986) or “weak hydrogen bonds” between aromatic rings and the π systems of the cyano groups (Meyer et al., 2003). However, we believe that in the present case due to the lack of proper directionality associated with classical hydrogen bond formation, the cyanovinyl-aromatic ring interaction is better described as a through-space electronic interaction. The conserved nature of the packing also highlights the preference of the molecule to make this type of contact. Similar interactions had been observed in the TMC278/RT wild-type and mutant structures, where cyanovinyl has been observed to participate in extensive interactions with a cluster of hydrophobic amino acids comprising the hydrophobic tunnel between NNIBP and external solvent (Das et al., 2008). Based on the geometry and alignment of the groups involved, we hypothesize that the nature of the cyanovinyl interaction with aromatic rings is highly favorable and allows for broad structural variability. Cyanovinyl moiety interactions in the pocket can be described as ‘slippery,’ permitting the drug molecule to slide up and down the hydrophobic tunnel while accommodating structural changes in the pocket of the protein.

Conclusions

The small molecule crystal structures of five DAPY NNRTIs presented in this work provide evidence of intrinsic flexibility of these molecules. Three crystal forms of TMC278 sample four alternative conformations. Hydrogen bonding of TMC278 to other symmetry related molecules and/or ordered solvents in the crystal resulted in changes of

bond distances at pyrimidine and linker regions. These observations provide an insight into possible conformation sampling in the context of RT HIV-1 binding pocket where the hydrogen bond network involves linker nitrogens of the drug. The structures suggest that besides the ability of the drug molecule to sample a wide conformational space via varying torsion angles, it is able to do so also by assuming alternative hybridization states at the linker nitrogens. The crystal packing of protonated TMC278 crystal form provides supportive evidence of amphiphilic nature of the molecule. Packing of molecules in the crystal and hydrogen bond network with ordered solvent suggest that the hydrophobic wings of the molecule behave analogously to the hydrophobic tail of a surfactant while a more polar pyrimidine ring and linker nitrogens correspond to a polar head group.

Materials and Methods

Three TMC278 crystal structures and one structure of TMC125 were solved in collaboration with Dr. T. Emge, Chemistry Department, Rutgers University. The three crystal forms of TMC278 included: protonated, bis-DMSO, and anhydrous. Protonated TMC278 form was crystallized using the vapor diffusion hanging drop method from 1 mM and 5 mM drug in 20 mM malonic acid, pH 1.5 with 6 μl drops over 500 μl reservoir. Drop content included 0.3 μl of 20 mM TMC278 in DMSO and 3 μl of 20 mM malonate buffer adjusted to pH 1.5 using 1 N HCl. Low drug concentration conditions produced mainly thin plates while higher concentration conditions had on average a larger number of thicker crystals (Figure 33, Table 2). TMC278 bis-DMSO crystals were

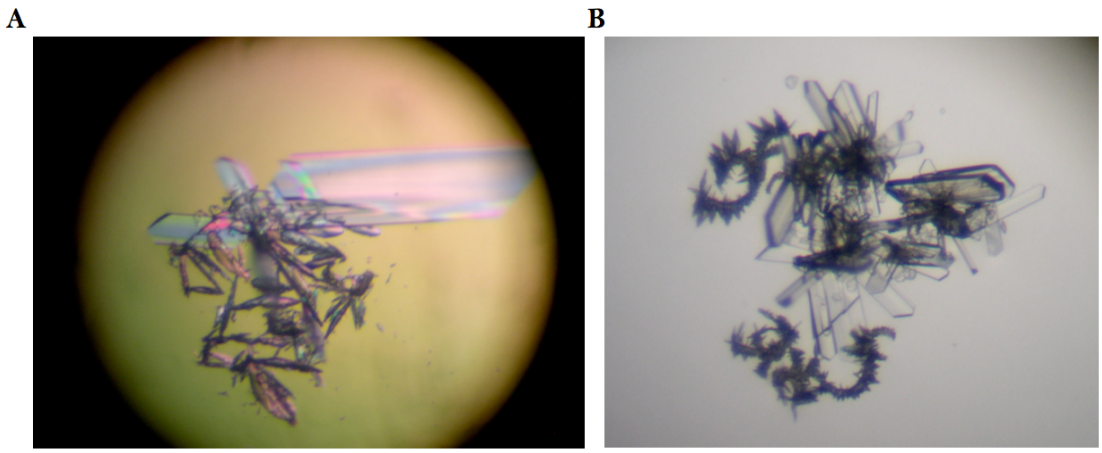


Figure 33. Photographs of TMC278 crystals grown using the hanging drop method in malonic acid, pH 1.5 at 1 mM **(A)** and 5 mM **(B)** drug concentrations.

obtained by slow evaporation method from DMSO, at starting concentration of 7.5 mM, over 24 hours. Anhydrous TMC278 crystals were re-crystallized from the bis-DMSO drops after initial crystals were fully dissolved and slow evaporation over an additional 48 to 72 hours was allowed. Anhydrous TMC125 crystals were also obtained by slow evaporation of a 5 mM drug concentration in DMSO over a 24-hour period.

Single-crystal X-ray diffraction data were collected for crystals on a Bruker Smart APEX CCD system. Data for the protonated TMC278 form was collected using a fine focus sealed tube with Cu Kα radiation while data on bis-DMSO, anhydrous TMC278, and TMC125 crystals was collected with fine focus sealed tube with Mo Kα radiation (see details in Table 2). Structures were solved by direct methods using SHELXS-97 (Sheldrick, 1997) and refined using SHELXL-97 (Sheldrick, 1997). Structure packing analysis was done using the Maestro graphical interface (Schrödinger, LLC).

Bond Angles [°] for TMC278.

	Protonated	bis-DMSO A	bis-DMSO B	Anhydrous
C(4)-N(1)-C(1)	117.74(19)	116.65(13)		116.49(11)
C(4)-N(2)-C(3)	120.1(2)	114.24(13)		113.57(11)
C(4)-N(2)-H(2N)	121.6(16)			
C(3)-N(2)-H(2N)	118.0(16)			
C(4)-N(3)-C(5)	127.03(19)	126.56(13)		130.19(11)
C(4)-N(3)-H(3N)	118.9(16)	114.6(14)		112.6(12)
C(5)-N(3)-H(3N)	113.9(16)	113.1(14)		116.5(12)
C(1)-N(4)-C(12)	123.68(19)	122.65(19)	122.5(4)	122.28(11)
C(1)-N(4)-H(4N)	115.7(16)		115.7(16)	114.3(12)
C(12)-N(4)-H(4N)	116.7(16)	117.8(15)	121.3(15)	118.1(12)
N(1)-C(1)-N(4)	117.5(2)	117.94(13)		117.88(11)
N(1)-C(1)-C(2)	122.46(19)	121.57(13)		121.68(12)
N(4)-C(1)-C(2)	120.1(2)	120.49(14)		120.44(12)
C(3)-C(2)-C(1)	116.8(2)	116.01(14)		115.52(12)
C(3)-C(2)-H(2)	121.6	122		122.3(11)
C(1)-C(2)-H(2)	121.6	122		122.1(11)
C(2)-C(3)-N(2)	120.4(2)	124.34(15)		125.12(12)
C(2)-C(3)-H(3)	119.8	117.8		118.0(11)
N(2)-C(3)-H(3)	119.8	117.8		116.9(11)
N(1)-C(4)-N(3)	120.2(2)	118.15(14)		118.64(11)
N(1)-C(4)-N(2)	122.5(2)	126.98(14)		127.42(12)
N(3)-C(4)-N(2)	117.34(19)	114.87(13)		113.94(11)
C(10)-C(5)-C(6)	119.6(2)	118.92(14)		117.96(11)
C(10)-C(5)-N(3)	117.4(2)	117.08(14)		116.81(11)
C(6)-C(5)-N(3)	123.0(2)	124.00(14)		125.23(11)
C(7)-C(6)-C(5)	120.2(2)	120.10(14)		120.37(12)
C(7)-C(6)-H(6)	119.9	120		118.5(10)
C(5)-C(6)-H(6)	119.9	120		121.1(10)
C(6)-C(7)-C(8)	120.0(2)	120.23(14)		120.78(12)
C(6)-C(7)-H(7)	120	119.9		118.7(10)
C(8)-C(7)-H(7)	120	119.9		120.6(10)
C(7)-C(8)-C(9)	119.9(2)	120.00(14)		119.17(12)
C(7)-C(8)-C(11)	119.6(2)	120.64(15)		121.11(12)
C(9)-C(8)-C(11)	120.5(2)	119.27(15)		119.72(12)
C(10)-C(9)-C(8)	120.0(2)	119.51(15)		119.98(12)
C(10)-C(9)-H(9)	120	120.2		120.0(11)
C(8)-C(9)-H(9)	120	120.2		120.0(11)
C(9)-C(10)-C(5)	120.3(2)	121.21(15)		121.66(12)
C(9)-C(10)-H(10)	119.9	119.4		120.9(10)
C(5)-C(10)-H(10)	119.9	119.4		117.4(10)
N(5)-C(11)-C(8)	179.2(3)	177.78(19)		178.73(15)
C(17)-C(12)-C(13)	122.4(2)	120.5(2)	120	121.25(12)
C(17)-C(12)-N(4)	117.5(2)	118.3(3)	118.5(4)	120.03(12)
C(13)-C(12)-N(4)	120.1(2)	121.1(3)	121.2(4)	118.71(12)
C(14)-C(13)-C(12)	117.8(2)	119.0(3)	120	118.95(12)
C(14)-C(13)-C(18)	119.7(2)	119.6(2)	118.9(5)	120.49(13)
C(12)-C(13)-C(18)	122.53(19)	121.3(3)	121.0(5)	120.54(13)
C(13)-C(14)-C(15)	121.6(2)	121.1(2)	120	121.24(12)
C(13)-C(14)-H(14)	119.2	119.5	120	118.2(11)
C(15)-C(14)-H(14)	119.2	119.5	120	120.5(11)
C(16)-C(15)-C(14)	118.5(2)	118.9(2)	120	118.54(12)
C(16)-C(15)-C(20)	119.1(2)	118.3(2)	122.0(4)	118.71(12)
C(14)-C(15)-C(20)	122.3(2)	122.8(2)	118.0(4)	122.74(12)
C(15)-C(16)-C(17)	121.9(2)	121.0(2)	120	121.82(12)
C(15)-C(16)-H(16)	119.1	119.5	120	117.6(9)
C(17)-C(16)-H(16)	119.1	119.5	120	120.5(9)
C(16)-C(17)-C(12)	117.8(2)	119.3(3)	120	118.11(12)
C(16)-C(17)-C(19)	120.5(2)	120.0(3)	120.0(5)	120.79(12)

Bond lengths [Å] for TMC278 (continued)

	Protonated	bis-DMSO A	bis-DMSO B	Anhydrous
C(20)-C(21)	1.328(3)	1.327(3)	1.319(6)	1.330(2)
C(20)-H(20)	0.95(2)	0.9500	0.9500	0.982(18)
C(21)-C(22)	1.433(3)	1.425(3)	1.431(7)	1.425(2)
C(21)-H(21)	0.96(3)	0.9500	0.9500	0.98(2)
O(1W)-H(1WA)	0.82(3)			
O(1W)-H(1WB)	0.88(4)			
S(1)-O(1)		1.5052(13)		
S(1)-C(23)		1.774(2)		
S(1)-C(24)		1.780(2)		
C(23)-H(23A)		0.9800		
C(23)-H(23B)		0.9800		
C(23)-H(23C)		0.9800		
C(24)-H(24A)		0.9800		
C(24)-H(24B)		0.9800		
C(24)-H(24C)		0.9800		
S(2A)-O(2A)		1.503(2)	1.511(10)	
S(2A)-C(26A)		1.782(2)	1.781(10)	
S(2A)-C(25A)		1.783(3)	1.790(10)	
C(25A)-H(25A)		0.9800	0.9800	
C(25A)-H(25B)		0.9800	0.9800	
C(25A)-H(25C)		0.9800	0.9800	
C(26A)-H(26A)		0.9800	0.9800	
C(26A)-H(26B)		0.9800	0.9800	
C(26A)-H(26C)		0.9800	0.9800	

Table 8-1. The full list of bond distances for TMC278 X-ray structures.

Hybridization analysis was performed using angle and distance measurements performed using SHELXS-97 (Sheldrick, 1997). Hydrogen atom deviation from the plane, called d_d, was determined by measuring the distance between the atom (HN3 or HN4) and the plane defined by three atoms (C5-N3-C4 or C1-N4-C12). The angle of displacement between the hydrogen atom and the corresponding plane, θ, was calculated using d_d and N…H bond distance (d_{hb}) in the following manner:

$$\theta = \arccos (d_d/d_{hb}) \qquad \text{EQ.1}$$

Crystal packing analysis consisted of four points: graphical inspection, ring-to-ring stacking distance, ring-to-ring angle, and total number of stacked rings. Graphical

Bond Angles [°] for TMC278 (cont.)

	Protonated	bis-DMSO A	bis-DMSO B	Anhydrous
H(18A)-C(18)-H(18B)	109.5	109.5		110.9(18)
C(13)-C(18)-H(18C)	109.5	109.5		114.7(12)
H(18A)-C(18)-H(18C)	109.5	109.5		104.6(17)
H(18B)-C(18)-H(18C)	109.5	109.5		104.6(18)
C(17)-C(19)-H(19A)	109.5	109.5		111.1(11)
C(17)-C(19)-H(19B)	109.5	109.5		112.3(12)
H(19A)-C(19)-H(19B)	109.5	109.5		110.5(17)
C(17)-C(19)-H(19C)	109.5	109.5		111.2(13)
H(19A)-C(19)-H(19C)	109.5	109.5		107.5(17)
H(19B)-C(19)-H(19C)	109.5	109.5		103.9(17)
C(21)-C(20)-C(15)	126.0(2)	125.9(2)	127.4(5)	127.17(13)
C(21)-C(20)-H(20)	118.8(13)	117.1	116.3	116.6(11)
C(15)-C(20)-H(20)	115.2(13)	117.1	116.3	116.2(11)
C(20)-C(21)-C(22)	122.6(2)	121.0(2)	121.5(5)	120.67(14)
C(20)-C(21)-H(21)	123.9(15)	119.5	119.2	122.0(12)
C(22)-C(21)-H(21)	113.4(15)	119.5	119.2	117.3(12)
N(6)-C(22)-C(21)	177.7(3)	179.2(3)	176.7(7)	178.41(16)
H(1WA)-O(1W)-H(1WB)	105(3)			
O(1)-S(1)-C(23)		106.37(9)		
O(1)-S(1)-C(24)		105.84(9)		
C(23)-S(1)-C(24)		97.58(11)		
S(1)-C(23)-H(23A)		109.5		
S(1)-C(23)-H(23B)		109.5		
H(23A)-C(23)-H(23B)		109.5		
S(1)-C(23)-H(23C)		109.5		
H(23A)-C(23)-H(23C)		109.5		
H(23B)-C(23)-H(23C)		109.5		
S(1)-C(24)-H(24A)		109.5		
S(1)-C(24)-H(24B)		109.5		
H(24A)-C(24)-H(24B)		109.5		
S(1)-C(24)-H(24C)		109.5		
H(24A)-C(24)-H(24C)		109.5		
H(24B)-C(24)-H(24C)		109.5		
O(2A)-S(2A)-C(26A)		106.38(12)		
O(2A)-S(2A)-C(25A)		106.90(15)		
C(26A)-S(2A)-C(25A)		96.83(13)		
O(2B)-S(2B)-C(26B)		104.5(12)		
O(2B)-S(2B)-C(25B)		103.9(12)		
C(26B)-S(2B)-C(25B)		96.0(10)		
S(2B)-C(25B)-H(25D)		109.5		
S(2B)-C(25B)-H(25E)		109.5		
H(25D)-C(25B)-H(25E)		109.5		
S(2B)-C(25B)-H(25F)		109.5		
H(25D)-C(25B)-H(25F)		109.5		
H(25E)-C(25B)-H(25F)		109.5		
S(2B)-C(26B)-H(26D)		109.5		
S(2B)-C(26B)-H(26E)		109.5		
H(26D)-C(26B)-H(26E)		109.5		

Table 8-2. The full list of bond angles in three TMC278 X-ray structures.

Hydrogen bonds for TMC278-Protonated [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(2)-H(2N)...O(1W)#1	0.84(3)	1.88(3)	2.721(3)	177(2)
N(3)-H(3N)...Cl(1)#2	0.88(3)	2.45(3)	3.294(2)	161(2)
N(4)-H(4N)...Cl(1)	0.82(3)	2.40(3)	3.228(2)	178(2)
O(1W)-H(1WA)...Cl(1)	0.82(3)	2.38(3)	3.182(2)	167(3)
O(1W)-H(1WB)...Cl(1)#3	0.88(4)	2.22(4)	3.084(2)	167(3)

Symmetry transformations used to generate equivalent atoms:
#1 -x+1,-y,-z+1 #2 x+1,y+1,z #3 -x,-y-1,-z+1

Hydrogen bonds for TMC278-bisDMSO [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(3)-H(3N)...N(2)#1	0.80(2)	2.23(2)	3.0289(19)	173.1(19)
N(4)-H(4N)...O(1)	0.81(2)	2.14(2)	2.9341(18)	168(2)

Symmetry transformations used to generate equivalent atoms:
#1 -x,-y+1,-z

Hydrogen bonds for TMC278-Anhydrous [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(3)-H(3N)...N(2)#1	0.88(2)	2.47(2)	3.337(2)	172(2)
N(4)-H(4N)...N(5)#2	0.86(2)	2.26(2)	3.089(2)	163(2)

Symmetry transformations used to generate equivalent atoms: #1 -x,-y,-z #2 x+1,-y-1/2,z-1/

Hydrogen bonds for R165 [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(4)-H(4NA)...N(2)#1	0.86(4)	2.24(4)	3.088(4)	166(3)
N(4)-H(4NB)...N(5)#2	0.93(4)	2.23(4)	3.056(4)	147(3)

Symmetry transformations used to generate equivalent atoms:
#1 -x+1,-y+2,-z+1 #2 x-2,y+1,z

Table 8-3. Hydrogen bonding statistics for TMC278 and TMC125 small molecule X-ray crystal structures.

analysis was conducted using Coot v0.1 applying corresponding crystallographic symmetry operators to generate a 15 Å nearest neighbor sphere around the reference molecule. Ring-to-ring distance (d_s) referred to the distance between two ring centers within 6 Å of each other (Figure 34A). Orientation of two rings participating in stacking interactions was evaluated using the ring-to-ring dihedral angle (<s). The dihedral angle measurement included both ring centers and two non-hydrogen ring atoms

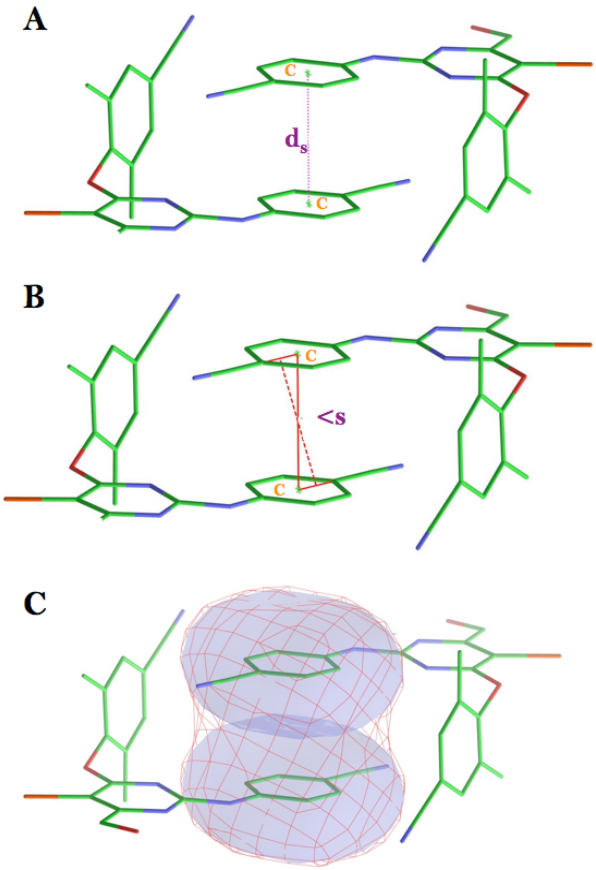


Figure 34. Stacking analysis method used in the small molecule crystal packing analysis. A. The stacking distance, d_s, between two rings is assessed by measuring the distance between two corresponding ring centers, C. B. Alignment of two stacking rings is evaluated by measuring dihedral angle, <s, describing rotation of rings relative to each other. C. Extent of the stacking interactions is assessed by evaluation of buried surface ($SESA_s$), where the difference between individual surface areas of two rings and the joined surface area is calculated (eq. 2).

(Figure 34B). The magnitude of stacking between two rings was evaluated using the buried surface area or solvent excluded surface area of stacking ($SESA_s$) (Figure 34C) calculated using the individual solvent accessible surface areas (SASA) of two stacking rings (R1 and R2) and a joined solvent accessible surface area of the stacking ($SASA_{total}$) using the following equation:

$$SESA_s = SASA_{total} - (SASA_{R1} + SASA_{R2}) \quad \text{EQ. 2}$$

Parallel stacking geometry was defined by d_s of 3.5 ± 0.5 Å , <s of either 180 ± 20° or 0 ± 20°, and $SESA_s \geq 20$ Å^2. Perpendicular ring-to-ring stacking was characterized by d_s of 3.5 ± 0.5 Å and <s of 125 ± 10°. All SESA were calculated using Maestro (Schrodinger).

Comparison between crystal packing modes was done using the quality of stacking (QS) parameter that was implemented to describe the combined contribution from d_s, <s, and $SESA_s$ descriptors. If all three parameters were falling within the above defined range the stacking interaction was classified as optimal and given 3 points. If either the distance between two rings or the angle or $SESA_s$ were falling outside of allowed range, the stacking interaction was classified as suboptimal and a penalty of 1 point for each was subtracted. If all three parameters were falling outside of the allowed range, the stacking interaction was considered to be not optimal. The total number of rings participating in intermolecular stacking interactions (N_{TS}) corresponds to the number of rings per individual drug molecule involved in stacking interactions.

Chapter 3. Amphiphilic behavior of hydrophobic drug molecules in MD simulation of TMC278 drug aggregate/water interface

Synopsis

Biophysical studies of NNRTI behavior in simple solutions mimicking physiologically relevant conditions established a correlation between ability of the molecules to form stable, homogeneous populations of clear, ~100 nm in diameter, spherical nanoparticles at low pH in surfactant independent fashion, and good oral bioavailability. We hypothesized that TMC278 is able to display surfactant-like properties under physiologically relevant conditions and therefore facilitate formation of nanostructures in the absence of other surfactants. We tested our hypothesis using all-atom molecular dynamics (MD) simulations with neutral and protonated drug molecules in explicit solvent. Results of MD simulations support our hypothesis and show that protonated TMC278 molecule at the aggregate/water interface behaves as a surfactant. We determined that drug aggregate surface in MD simulation is populated by both neutral and protonated molecules. Protonated TMC278 molecules behave as amphiphiles at the aggregate/water interface by displaying hydrophilic and burying hydrophobic portions of the molecule while neutral TMC278 molecules sample large conformational variability and enhance aggregate packing. Our simulation results suggest that intrinsic torsional flexibility and ability to exhibit amphiphilic properties at low pH facilitate formation of a stable aggregate/water interface in drug aggregates and therefore are thought to play critical roles in oral bioavailability properties of the drug.

Introduction

Biophysical measurements, reported in the first chapter, confirmed formation of aggregate nanostructures and showed that oral bioavailability properties of the DAPY compounds correlated with their ability to form homogeneous populations of small (<120 nm diameter) aggregates at low pH and high drug concentration (Frenkel et al., 2005). Concurrent correlation analysis for the DAPY NNRTI molecules using 15 physicochemical, pharmacokinetic, and chemical parameters had shown strong correlation between aggregate formation, lipophilicity, transepithelial transport, and human exposure (Lewi et al., 2004). As an outcome of these analyses, we formulated a hypothesis that small drug aggregates were the biological units of delivery that formed at high drug concentrations in the stomach and were being further trafficked into lymphatic circulation by small intestine enterocytes via either paracellular or transcellular routes (Frenkel et al., 2005). Based on experiments reported by Frenkel et al., 2005 and those presented in this chapter, we conclude that ability to form small drug aggregates by TMC278 is independent of a stabilizing agent, therefore, suggesting that TMC278 along with few other members of NNRTI DAPY series is able to achieve nanoparticle formation purely due to its physicochemical characteristics. To test this hypothesis, we designed a series of computational experiments.

Among currently known nanoparticulate structure types, we have reasons to believe that microemulsion is the most appropriate model to describe NNRTI drug aggregates. Microemulsions are isotropic, thermodynamically stable ternary or pseudoternary systems with 80-100 nm diameters, otherwise known as swollen micelles that consist of oil (hydrophobic phase), water (hydrophilic phase), surfactant (stabilizing

agent), and frequently a cosolvent (de Gennes and Taupin, 1982; Eicke and Parfitt, 1987; Rosen, 1989; Tsujii, 1997). Surfactants are known to lower the surface tension between two immiscible phases (de Gennes and Taupin, 1982; Eicke and Parfitt, 1987; Rosen, 1989; Tsujii, 1997) and to have specific structural characteristics such as charged or polar head groups and hydrophobic tails. In the case of some DAPY NNRTIs we hypothesize that under acidic conditions, protonated drug molecules have surfactant-like features with the protonated pyrimidine ring being a head group and the diaryl wings composing the hydrophobic tail. We propose that those NNRTI DAPY compounds that can exhibit surfactant-like properties upon protonation can facilitate microemulsion-like aggregate formation.

Here, TMC278 behavior at the water/aggregate interface was examined using all-atom molecular dynamics (MD) simulations. A number of computational studies have been targeted at water/vapor interfaces (Chen, 1995) and microemulsions such as simulations of single or multiple surfactant molecules in explicit solvent (Derecskei et al., 1999), coarse-grained simulations using self-consistent field theory (Fraaije et al., 1997), and dissipative particle dynamics (DPD) based on nuclei formation and growth via coalescence-exchange or coagulation of water-in-oil drops for simulation of microemulsions, micelles, etc. (Chen et al., 2007; Ryjkina et al., 2002) However, all-atom simulations of aggregate systems consisting of a half million molecules or more is not currently feasible. Therefore, a computational approach was devised to study the surface dynamics of an aggregate rather than the full nm-sized aggregates. In the present study, pH-dependent aggregate surface formation was monitored with the goal to

evaluate the behavior of protonated and neutral drug molecules in terms of their potential surfactant-like behavior.

Results

Infinite dilution studies of TMC278 and TMC125 in water

Here we report results of a computational study exploring conformational sampling for two NNRTI drugs, TMC278 and TMC125, under minimally constrained conditions of an infinite dilution in water. Previous structural studies, as well as computational studies reported in the subsequent sections of this work, focus on the NNRTIs' dynamics in the context of either a binding pocket, an aggregate, or a crystal lattice, therefore, asking question of behavior of the drugs under constrained conditions. Yet, the unconstrained behavior of the drug molecules is not well explored in any of these systems. Understanding of the molecular dynamics in solution serves the purpose of setting a reference point for analysis of NNRTI dynamics at the water/aggregate interface.

Conformational variability of TMC278 and TMC125 in water

Analysis of infinite dilution simulation results for DAPY NNRTIs suggests that they sample a substantially larger conformational landscape in water than in the binding pocket of HIV-1 RT. As can be seen from Figure 35, we analyzed torsional flexibility of the molecules in MD simulation by monitoring changes of the dihedral angles $\tau 1 - \tau 4$.

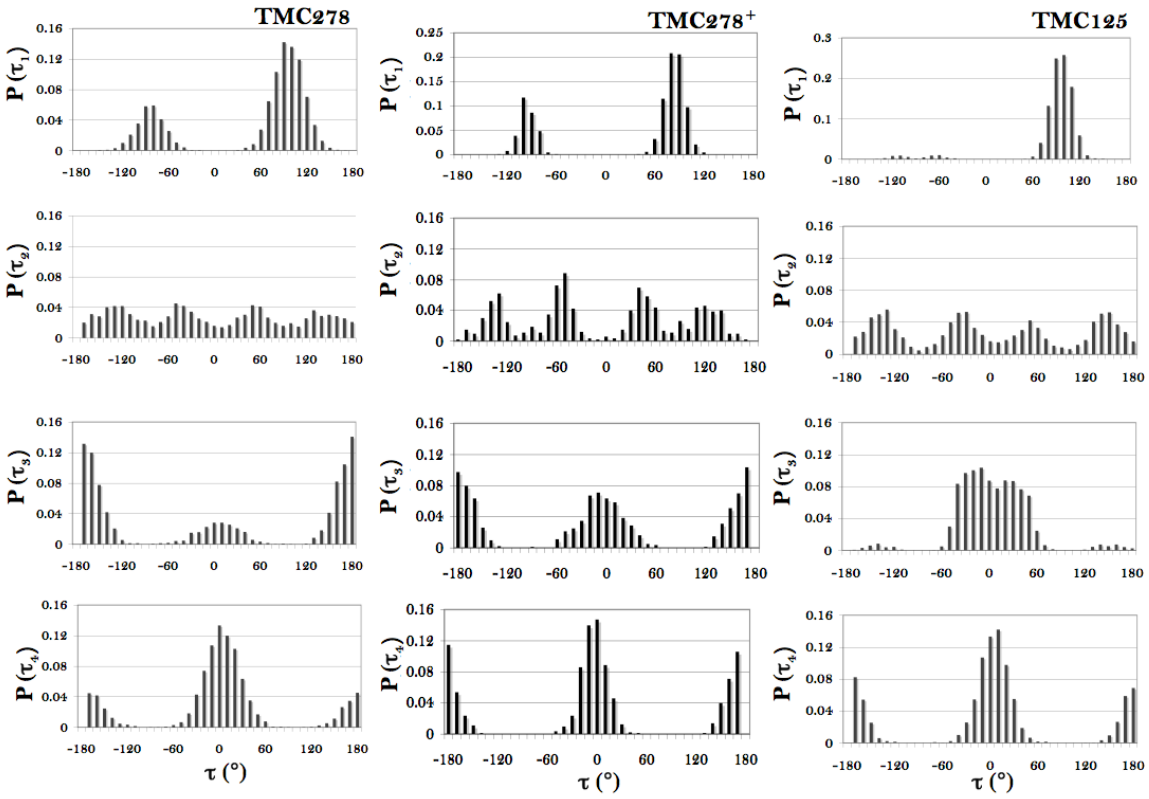


Figure 35. The torsion angle probability distributions for TMC278 in neutral and protonated forms and neutral TMC125 in the 8 ns infinite dilution condition REMD simulations, at 300 K.

Torsion angle distributions for neutral and protonated forms of TMC278 are rather similar. Both show earlier predicted broad distributions at angles τ1 and τ2, though τ2 displays substantially more rotational freedom than τ1 probably due to the presence of methyl groups on the phenyl ring that restrict the full rotation of the wing I moiety. Besides high rotational variability at the τ1 and τ2 angles, we also observed broad rotations of τ3 and/or τ4 torsion angles that have not been yet seen for TMC278 or TMC125 either in the binding pocket of the target protein or in the small molecule structures. Bound and small molecule drug conformations had shown 10-15° variations in τ3 and τ4, yet in the infinite dilutions results conformations are observed that are

rotated by 180°. These rotations change dramatically the overall shape of the molecule, resulting in more extended, "seahorse-like" conformations. TMC125 had shown to be more conservative in its overall conformational sampling than TMC278. Though rotation of wing II via τ2 and τ4 were similar for two drug molecules, the sampling of the broad rotations via dihedral angles τ1 and τ3 was not observed. In TMC125, an unfavorable interaction between aromatic ring and bromine group on the pyrimidine ring that results from a 180° rotation at τ3 were responsible for a low occurrence of such conformation. Based on these observations, the conformational landscape of the TMC125 is concluded to be more restricted than that of TMC278 in neutral and protonated forms.

In order to study behavior of DAPY NNRTIs in the molecular dynamics simulations we have categorized observed drug conformations into four main groups. Drug conformations with τ3 and τ4 less than 90º were determined to be folded, "U" conformations (Figure 40) and had been previously observed in crystallographic structures of the drug in protein complexes (wild-type and mutant)(Das et al., 2008), as well as in the small-molecule crystal structures (Frenkel et al., in preparation). Drug conformations with τ3 greater and τ4 less than 90º result in rotation of wing I relative to pyrimidine ring and were called flipped $L_{\tau 3}$, while those with τ3 less and τ4 greater than 90º, where wing II was rotated relative to pyrimidine ring, were called flipped $L_{\tau 4}$ (Figure 40). The most torsionally flexed conformation of TMC278 was called "E," for extended, and was characterized by τ3 and τ4 being greater than 90º, where both wings were rotated relative to pyrimidine (Figure 40). The E conformation has been previously observed for a diaryltriazine analog of a DAPY NNRTI molecule, R120393, in complex

with HIV-1 RT and in a small-molecule crystal structure (Das et al., 2004; Frenkel et al., in preparation).

Based on biophysical data presented in the first chapter it was hypothesized that protonation of the drug molecules changes their dynamics in solutions. We have explored charge-dependent behavior in the infinite dilution simulations by comparing protonated and neutral TMC278 MD simulations. Under neutral conditions, the molecule experiences more of the flipped or extended conformations in the simulation by rotating $\tau 3$ than in the commonly observed folded conformation where both $\tau 3$ and $\tau 4$ are less than 90° (Figure 35). Under protonated conditions the occurrence of fully rotated $\tau 3$ angles is estimated to be approximately 50% less, therefore suggesting that protonated molecules prefer folded conformations to that of flipped or extended in the MD simulation. In difference to the effects observed in the case of the $\tau 3$ sampling, $\tau 4$ conformation occurrence frequency did not change significantly with protonation at the pyrimidine ring. In fact, a slight increase (by ~2%) was observed (Figures B4).

Common shape conformation sampling frequencies were compared for all three infinite dilution simulations and determined that both protonation and unfavorable intramolecular interaction affect the conformational landscape. Sampling probability distributions were calculated for both simulations and are reported in Table 9. Sampling of the folded, U state, was more prominent in the protonated simulation with a probability of 0.43 versus in the neutral experiment where U conformation had sampling probability of only 0.06. The highest probability of folded state sampling was determined for TMC125, 0.59. Under neutral conditions, TMC278 molecule spent most of the time in the $L_{\tau 3}$ conformation while as protonated TMC278 did not favor partially extended

conformations. In the case of TMC125, for the partially extended conformations only $L_{\tau 4}$ was observed (probability 0.34). Surprisingly, the fully extended conformations of TMC278 were observed at comparatively high probability rates of 0.42 and 0.44 under

Shape Probability	U	$L_{\tau 3}$	$L_{\tau 4}$	E
TMC278 Neutral	0.06	0.49	0.03	0.42
TMC278 Protonated	0.43	0.01	0.12	0.44
TMC125 Neutral	0.59	0.06	0.34	0.02

Table 9. Common shape sampling probability in TMC278 and TMC125 REMD simulations at 300K.

both neutral and protonated conditions respectively, while as TMC125 was mostly unable to sample these conformations.

Energetics of neutral and protonated TMC278 in water

In addition to looking at the conformational changes observed by the drug molecules in water, their corresponding energies in both protonated and neutral infinite dilution TMC278 simulations were compared. As shown in Table 10, we compared potential energies of four common molecule shape groups of TMC278 in both neutral and protonated forms. In general protonated TMC278 molecules had shown lower average potential energy than the neutral form of the drug. Among different shape forms one the most energetically favorable was observed to be E conformations with -108 kcal/mol in neutral form and -133 kcal/mol in protonated. Of the four sampled conformations the least energetically favorable was folded, U conformation with -105.0

Average PE (kcal/mol)	U	L_{t3}	L_{t4}	E
Neutral	-105.0	-108.2	-107.4	-108.0
Protonated	-119.3	-127.6	-129.4	-133.1

Table 10. Average potential energy for neutral and protonated TMC278 in REMD simulation, at 300K.

and -119.3 kcal/mol in neutral and protonated systems respectively. Overall the difference in energies between different shape groups in the neutral system is ~2 kcal/mol, while as in the protonated simulation the average difference is approximately 7 kcal/mol. To determine the source of difference between energetic descriptors in the simulations, the individual energy component contributions were analyzed. Figure 36 presents the energy distribution plots for six energy components contributing to the total potential energy (PE): stretching, bending, torsional, van der Waals, electrostatic, and solvation energies. As can be seen from Figure 36, the main differences in energy of various molecular shape groups of TMC278 are in electrostatic and solvation energies, while stretching, bending, torsion, and van der Waals energies remain mainly unchanged. Protonation of the pyrimidine ring in TMC278 clearly changes the electrostatic interactions in the molecule and as a consequence affects the water molecule ordering in the region. In MD simulations, favorable electrostatic energies were observed in the

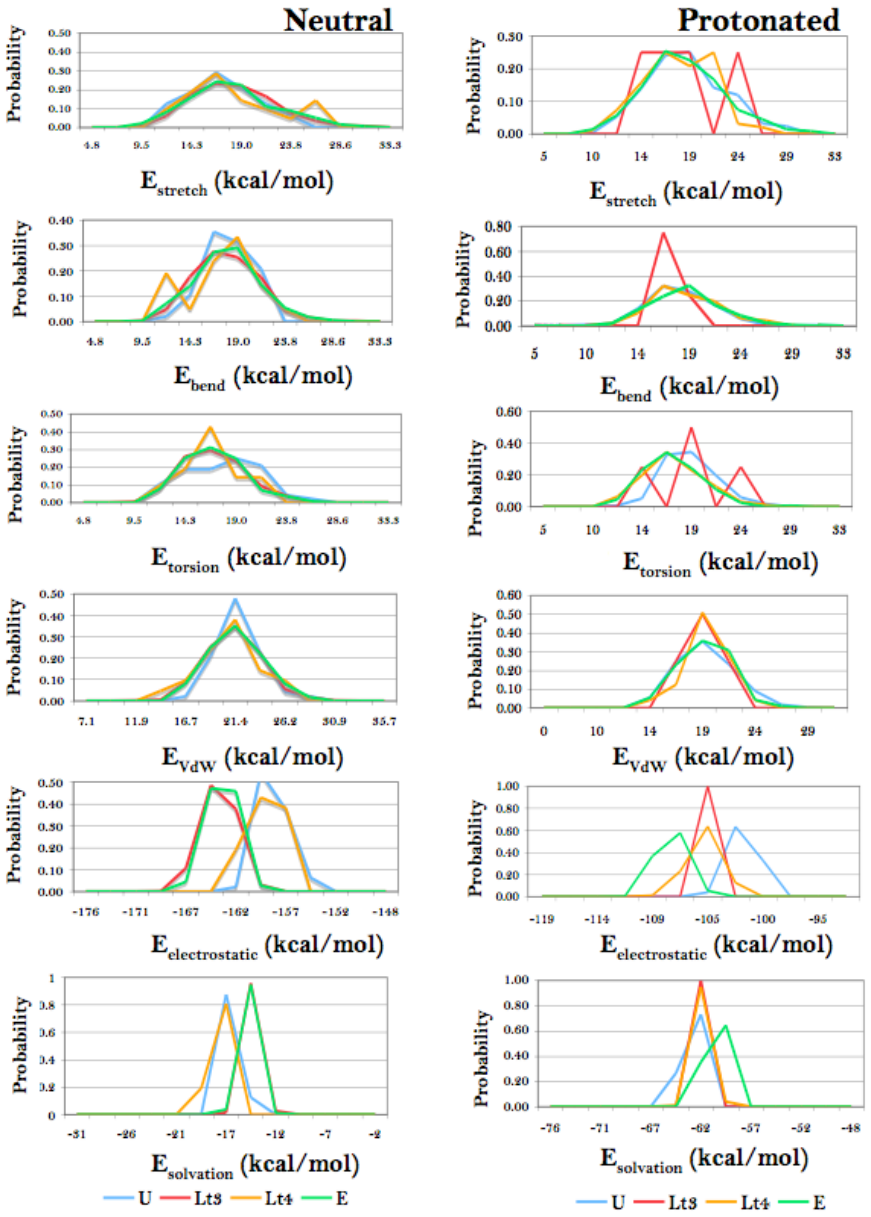


Figure 36. Potential energy component distribution plots for neutral and protonated TMC278 in the 8 ns infinite dilution REMD simulations. **A.** Neutral TMC278 potential energy distributions for U, $L_{\tau}3$, $L_{\tau}4$ molecular shape classes. Dashed lines correspond to median energy values for U (blue) and $L_{\tau}3$ (red) populations. **B.** Protonated TMC278 potential energy distribution for U, $L_{\tau}3$, and $L_{\tau}4$ molecular shapes. Dashed lines correspond to median potential energy values for U (blue) and $L_{\tau}4$ (yellow-green) populations.

neutral TMC278 model, while solvation energy was more favorable in the protonated drug form (Figure 36). In summary, infinite dilution experiment results suggest that protonated NNRTI drug molecules have different conformational sampling tendencies than their neutral analogs. Of the two different forms of TMC278, the neutral molecule had the most conformational freedom. Although behavior of protonated and neutral TMC278 was observed to differ in their corresponding preference for the folded conformation, both showed surprising preference for the extended conformation. One possible explanation behind the high preference of sampling of the fully extended conformations is that hydrophobic interactions between wings I and II in the folded state are being compensated by the intramolecular electrostatic/hydrophobic interaction and drug-water interactions.

Computational studies of TMC278 aggregates

Based on *in vitro* observations discussed in the first chapter, we conclude that TMC278 aggregation does not depend on the presence of added surfactant only at low pH conditions. Based on the TMC278 drug pKa of 5.6 and observed *in vitro* aggregation tendencies, we hypothesize that the surface-active agent responsible for the formation of the aggregates in the absence of tyloxapol is the protonated (positively charged) form of TMC278. The tendency towards larger aggregates with increasing pH is thus rationalized in terms of the decrease of the fraction of surface-active protonated TMC278 molecules. To understand further the effect of protonation on the drug aggregate assembly processes in aqueous environments we conducted a series of MD simulations to further probe the protonation-dependent behavior of the drug.

To validate the hypothesis stated above and to unambiguously investigate the pH-dependent aggregation properties of TMC278, we have formulated minimalist models of solutions containing only water, TMC278 molecules (of which 0 to 90% were charged),

	0%	10%	30%	60%	90%
Computational Models					
Total # of TMC278 Molecules	100	100	100	100	100
# of Protonated Molecules	0	10	30	60	90
# of Cl^- Atoms	0	10	30	60	90
# of Water Molecules	5135	4783	4720	4582	4324
Final Aggregate Composition*					
N_{agg}	98	98	85	46	0**
# of Surface Molecules	48	48	49	36	0**
# of Core Molecules	50	50	36	10	0**
# of Protonated Surface Molecules	0	9	17	3	0**
SASA ($Å^2$)	10610	11663	10728	6950	0**
Surface Charge Density (e/nm^2)	-	0.008	0.016	0.004	0**
U Conformation Populations***					
Aggregate Core	99.8	65.3	61.8	-	-
Aggregate Surface, Neutral	90.9	60.3	45.6	-	-
Aggregate Surface, Protonated	-	63.5	58.9	-	-
$L_{\tau 3}$ Conformation Populations***					
Aggregate Core	0.2	27.9	37.2	-	-
Aggregate Surface, Neutral	9.1	35.9	51.1	-	-
Aggregate Surface, Protonated	-	28.3	11.2	-	-

* aggregate composition was determined using the solvation parameter S_m (see Methods for details)
** not applicable due to lack of aggregation
*** fraction in percent

Table 11. Composition and analysis of the TMC278 aggregate MD results.

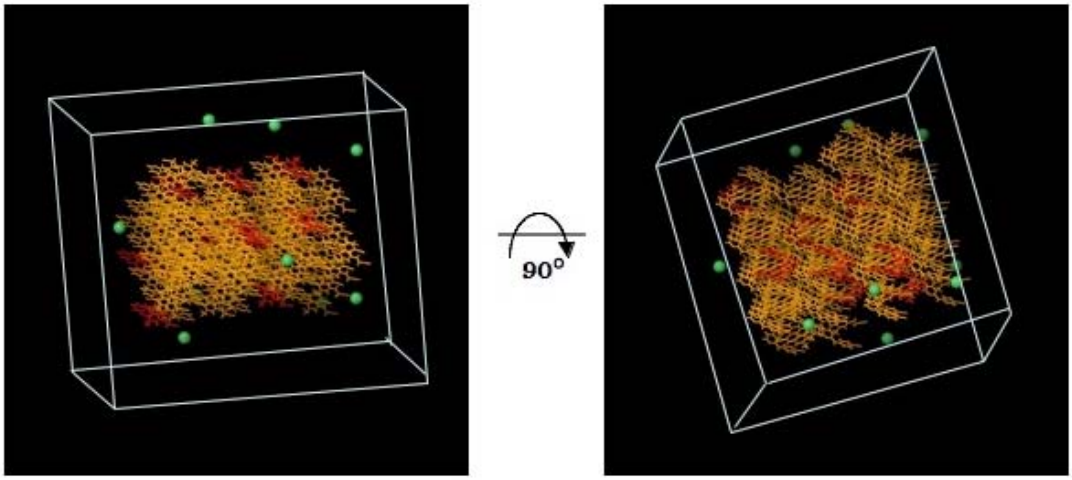


Figure **37.** The ten percent starting model in a 60 x 60 x 60 Å cell. Color assignment: yellow: neutral TMC278; red: charged TMC278; green: chloride ion. SPC waters are not shown.

and sufficient chloride ions to neutralize the system (Table 11, Figure 37). Morphology, charge distribution, and packing analysis suggest that drug aggregate structure is pH dependent and biphasic with a partially neutral core and a more disordered and partially charged surface.

Visual inspection of the aggregate morphology during the MD simulations of the TMC278/water systems at the increasing level of protonation indicates that aggregate formation is a pH-dependent phenomenon. The original cubic shape of the microcrystal (Figure 38) is lost for all but the model without protonated TMC278 molecules (0% protonation). The models with 10 and 30% protonated TMC278 molecules produce a continuous aggregate phase while those with 60 and 90% protonation results in either partially or fully dissociated aggregates. As reported in Figure 38 and Table 11, the 0, 10, and 30% protonated models yield structures with similar aggregation numbers (N_{agg}) and solvent-accessible surface areas (SASA), whereas, the 60% and 90% protonated

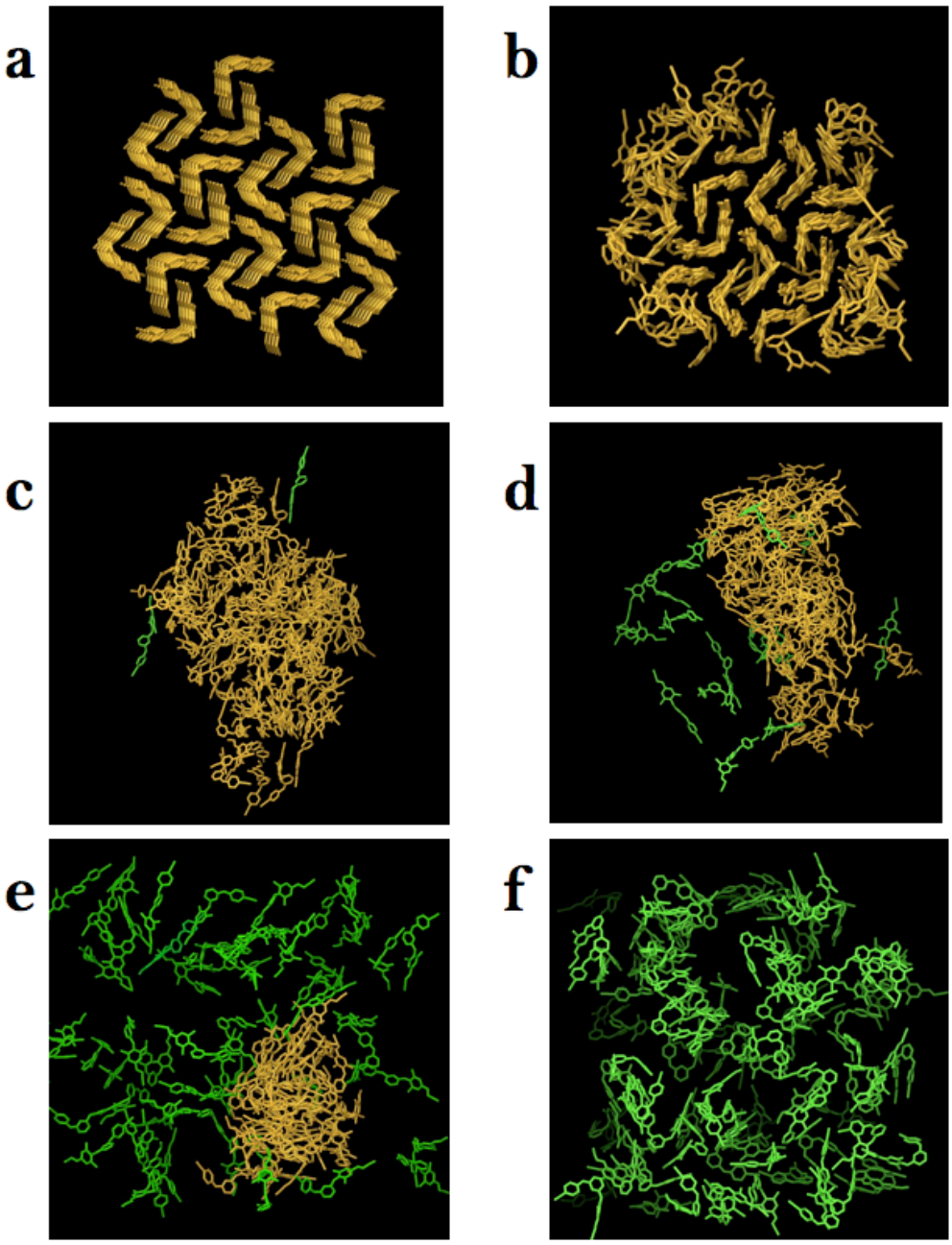


Figure **38.** Snapshots of TMC278 aggregate structures extracted from the MD trajectories at various levels of protonation. **a.** Starting conformation of the 0% protonation model. **b-f**. Final conformations of the 0, 10, 30, 60, and 90% protonation models. Water molecules and counter ions are not shown. Color assignment: yellow: aggregated TMC278 molecules; green: dissociated TMC278 molecules (aggregate association is measured based on S_m parameter described in Methods.)

models produced structures with significantly smaller N_{agg} and SASA values and therefore were not used in quantitative comparison analysis.

Molecular composition analysis of the aggregates reported in Table 11 indicates that charge saturation of the aggregate surface is reached at under 30% protonation. While the number of drug molecules making up the surface of the aggregates in the 10 and 30% protonated models is similar (48 and 49 molecules, respectively) the corresponding number of charged molecules at the surface are different (9 and 17, respectively). Consequently we observe that a three-fold increase in number of charged molecules (from 10 to 30%) is accompanied by only a two-fold increase in the surface charge density, from 0.008 to 0.016 charge/nm^2 respectively. The non-linear trend in surface charge distribution suggests that system reaches surface charge saturation at under 30% protonation. Furthermore, full dissolution of the aggregate at 90% protonation (Figure 38f) after only two nanoseconds of MD simulation suggests that under sufficiently acidic conditions aggregate formation would be inhibited.

Visual examination and quantitative packing analysis of the structures obtained from the simulations indicate that the aggregates tend to assume biphasic formation. The structures of the aggregates obtained at 0% protonation represent extreme examples of this phenomenon. In Figure 38b, we see that the aggregate displays two clearly distinguishable layers, one of which has retained the starting crystalline packing and the other, at the water interface, is partially disordered. Interestingly, such clear differentiation is present only at 0% protonation. Even at only 10% protonation (Figure 38c), the core of the aggregate does not display crystalline ordering.

Our results suggest that protonated TMC278 molecules tend to induce a liquid-like disorder at the water/aggregate interface that in turn likely induces disorder in the core of the aggregates. The water/aggregate interface of the 10% and 30% protonated models is populated preferentially by protonated molecules and displays greater disorder than the interface of the 0% model. The natural implication of this observation is that crystalline-like ordering would be more prevalent in regions deep in the aggregate core distant from the surface layer if larger number of drug molecules would be employed in atomistic simulations. This hypothesis is strengthened by the resiliency displayed by the crystalline-like ordered phase at 0% protonation, which, unlike at higher levels of protonation, remained stable for the entire duration of the simulation. In conclusion, although the modeling design employed here is not intended for studying the core structure of the aggregates, the present *in silico* results indicate that the nm-sized TMC278 aggregates we observed *in vitro* likely contain cores largely unaffected by the specific solution pH conditions, which are instead predicted to have a marked effect on the structure of the water/aggregate interface.

Ring-to-ring distance distribution analysis of the MD trajectories confirms that TMC278 molecules form a liquid-like layer at the water/aggregate interface while the core molecules of the 0% model preserve the crystalline arrangement. Distance analysis shows that stacking interactions characteristic of crystal packing are primarily present only in the 0% protonated model and are sampled to a lesser extent in the higher protonation systems (Figure 39a-b). Stacking between wing II and the pyrimidine ring is clearly present in all simulations and is hypothesized to be relevant to the drug aggregate surface structure at the water interface. Analysis of ring-to-ring distance distributions at

0% protonation for core and surface molecules confirms that the internal structure of the aggregate displays crystal ordering while the surface is more disordered or liquid-like (Figure 39d).

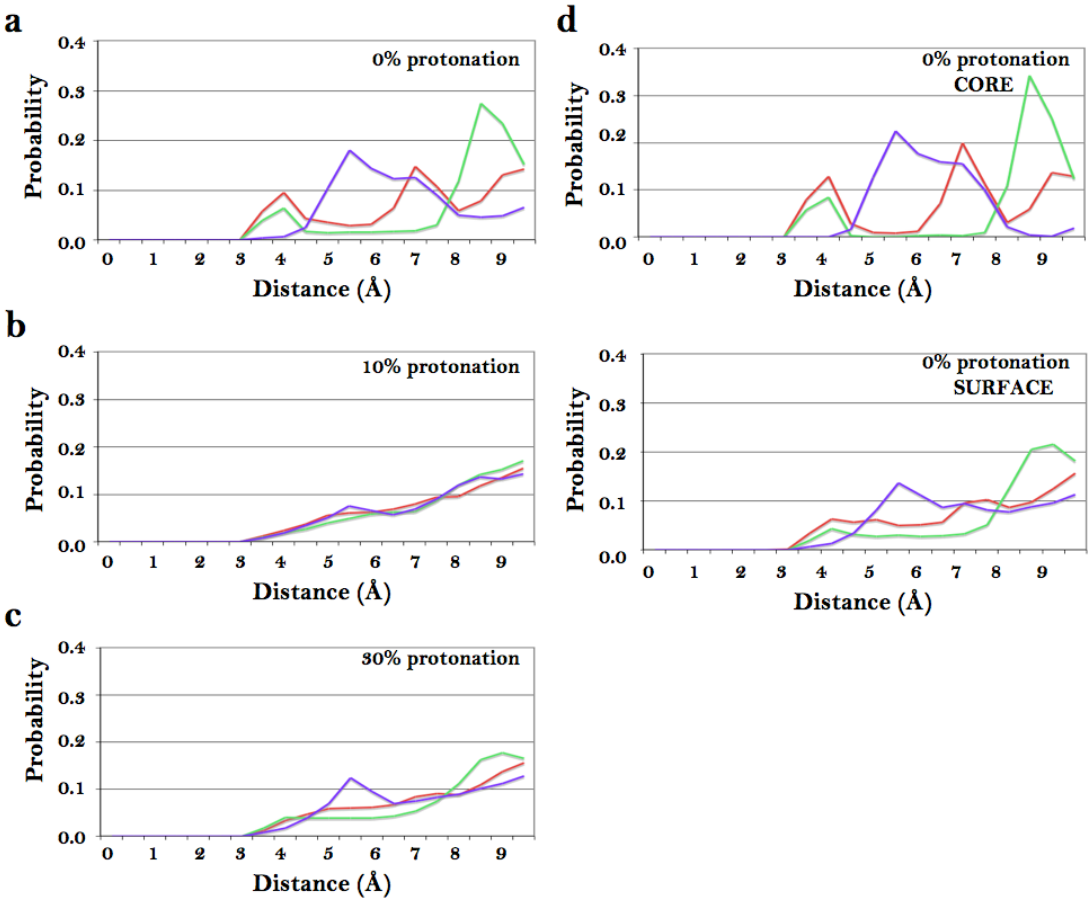


Figure 39. Intermolecular ring-to-ring distance distributions for the 0, 10, and 30% protonation models. **a.** 0%; **b.** 10%; **c.** 30%. **d.** Ring-to-ring distance distribution functions for the core TMC278 molecules and surface TMC278 molecules from the 0% protonation model. Color assignment: red: distances between ring I and ring II; green: distances between ring II and ring II; blue: distances between ring II and pyrimidine ring.

Surface properties of TMC278 aggregate models

Conformational and surface area analysis of TMC278 molecules at the aggregate-water interface suggests that torsional flexibility allows TMC278 to adopt surfactant-like orientations by exposing preferentially their "polar head" and burying their "hydrophobic tail". We have categorized observed drug conformations in MD trajectories into four main groups: U, $L_{\tau 3}$, $L_{\tau 4}$, and E (see Methods for details, Figure 40). The probability

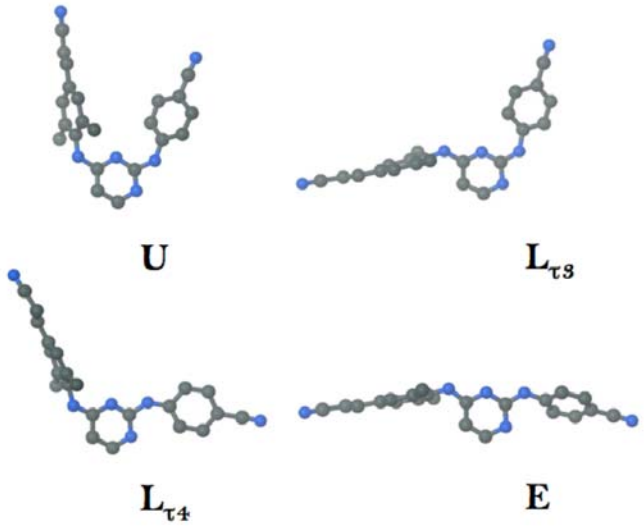


Figure 40. Representative conformations of TMC278 observed in simulations: U conformation ($\tau 3$ and $\tau 4$ values between -90º and 90º); $L_{\tau 3}$ conformation (-90º < $\tau 4$ < 90º and 90º < $\tau 3$ <-90º); $L_{\tau 4}$ conformation (-90º < $\tau 3$ <90º and 90º < $\tau 4$ >-90º); E conformation ($\tau 3$ >90ºand $\tau 4$ <-90º).

distributions for the $\tau 3$ and $\tau 4$ dihedral angles for core, surface, and dissociated molecules were used to monitor conformational sampling in MD simulations (Figure 41) and consistently indicate that the occurrence of partially and/or fully extended TMC278 conformations ($L_{\tau 3}$, $L_{\tau 4}$, and E) correlates with a higher fraction of protonated drug molecules in the system. U conformations are predominant in the core of the aggregates in all simulations. Neutral TMC278 molecules at the surface in 10 and 30% protonated

models show preference for $L_{\tau 3}$ conformations (35.9 and 51.1% probability, respectively, see Table 11), while protonated TMC278 molecules at the surface of the aggregates prefer the U conformation with probability of 63.5 and 58.9%, respectively. Therefore,

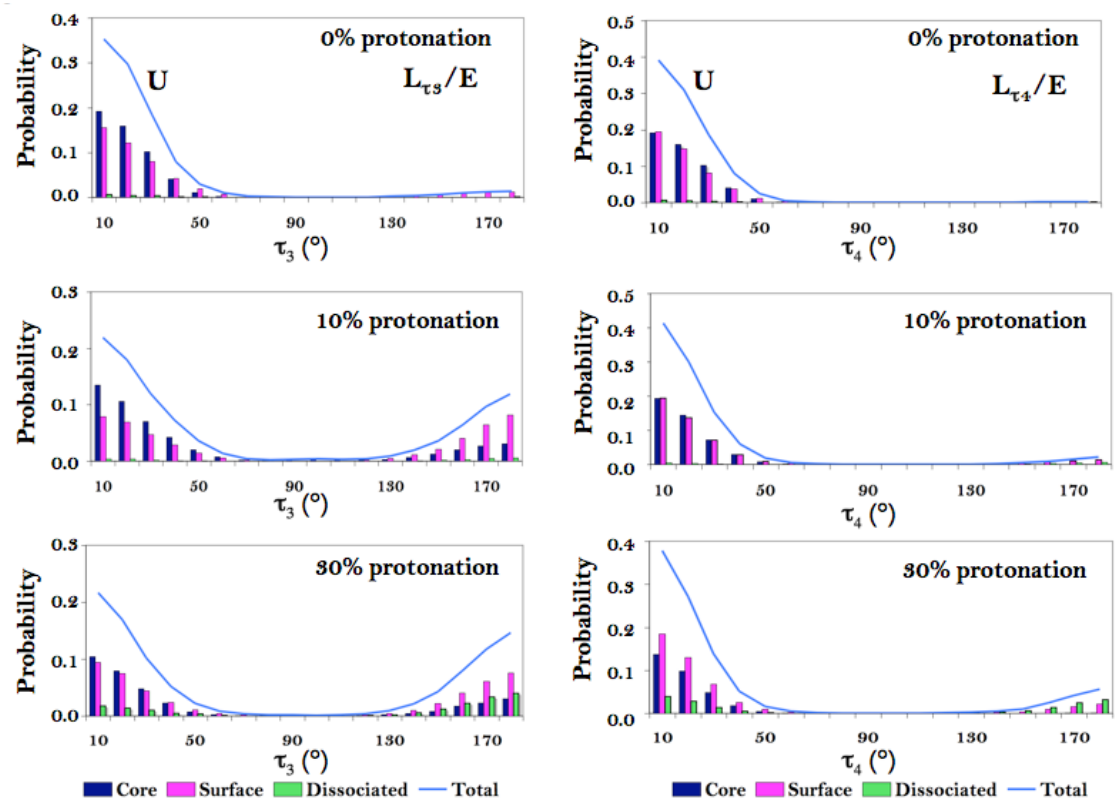


Figure 41. Occurrence of different TMC278 conformations in 0, 10, and 30% protonated model simulations. Probability distributions of TMC278 torsional angles τ3 (A) and τ4 (B) from the 0, 10, and 30% protonation models. Color assignment: light blue: all molecules in the simulation; dark blue: core molecules; magenta: surface molecules; green: dissociated molecules.

we conclude that protonated TMC278 molecules disfavor sampling of the $L_{\tau 3}/L_{\tau 4}$/E conformations while neutral molecules exist in a broader range of conformations. These observations highlight differences between the behavior of neutral and charged molecules on the aggregate surface.

In addition to looking at the drug aggregate packing motifs and conformational preferences, we also evaluated the makeup of the aggregate surface in terms of individual TMC278 atoms. Atom hydration analysis based on average normalized SASA of each heavy atom (see Methods, Figure 42) show, as expected, that core TMC278 molecules contribute little to the aggregate surface; however those core molecules that present atoms at the surface tend to display either cyano groups or pyrimidine rings. On the other hand, surface molecules, both neutral and protonated, prefer to expose wing II and pyrimidine rings. Hydration of neutral molecules at the surface is fairly uniform (Figure 42) with preferences towards exposure of the pyrimidine ring for the 0% model and the cyano group in the 30% model. Charged surface molecules preferentially expose their

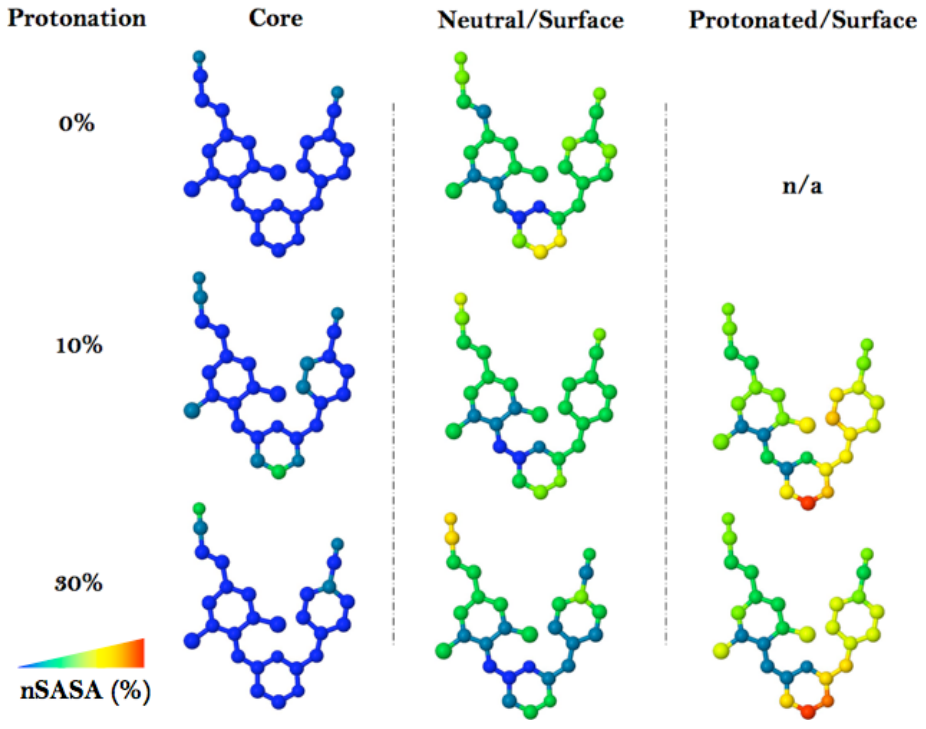


Figure 42. TMC278 surface contribution analysis. TMC278 atomic average solvent-accessible surface contributions computed from the 0, 10, and 30% protonation model MD trajectories. Color assignment is based on normalized SASA values ranging from minimum (blue) to maximum (red) hydration.

pyrimidine rings at the surface of the aggregate with greatest hydration of the C(3) atom, shown in red (Figure 42). One example of such surface exposure is presented in Figure 43 where protonated molecules in the 10% models are shown to expose preferentially pyrimidine ring at the water/aggregate interface. Based on these results we conclude that the contribution of core molecules to the surface is largely limited to cyano groups while surface molecules mostly display cyano groups, wing II and pyrimidine rings. These

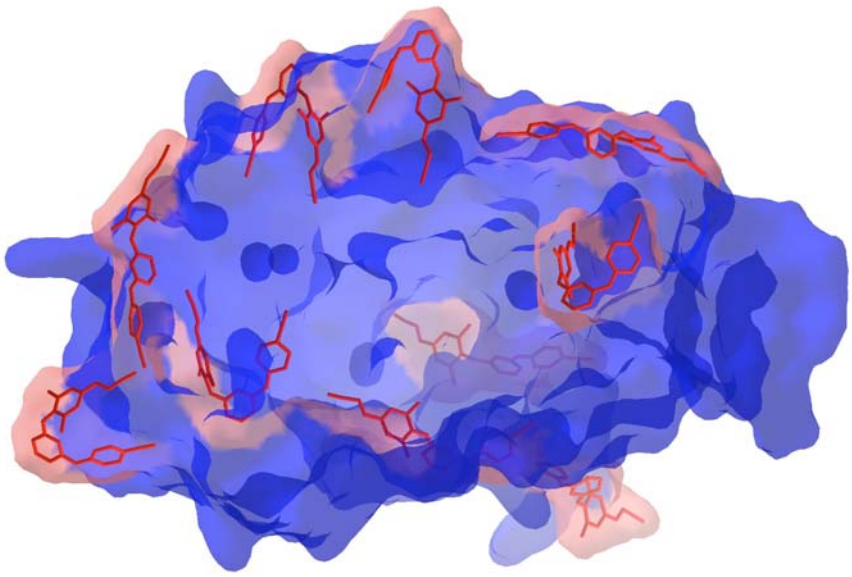


Figure 43. Conformational sampling by protonated TMC278 in 10% protonation model. Color assignment: red: protonated: blue: neutral molecules.

findings agree with modes of binding determined in structural studies of RT/TMC278 complexes, where it has been observed that the cyanovinyl group of the molecule is oriented towards the solvent through a hydrophobic tunnel while the pyrimidine ring is partially exposed to the solvent region and hydrophilic residues on the other side of the binding pocket (Das et al., 2008).

Taken together, these observations suggest that TMC278 molecules can exhibit different conformational preferences based on their environment. In Figure 44 we summarize our conclusions concerning the orientations of the TMC278 molecules at the water/aggregate interface. Core molecules in the aggregate preferred the U conformations while neutral, surface molecules sampled multiple conformations with

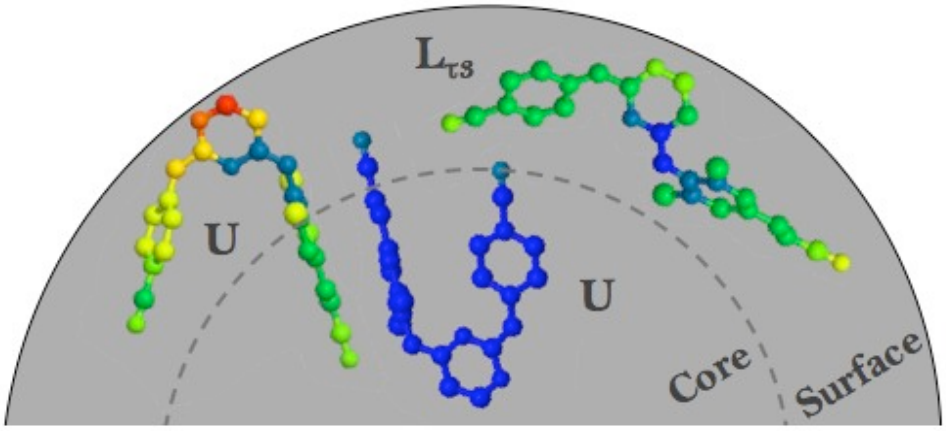


Figure 44. TMC278 surface contribution summary. **S**chematic diagram summarizing the observed preferred conformations and corresponding location (core and surface) in the aggregate. The diagram shows, starting from the left, a protonated molecules at the surface of the aggregate, a neutral molecule in the core of the aggregate, and a neutral molecule at the surface of the aggregate. Color assignment is based on normalized SASA values ranging from minimum (blue) to maximum (red) hydration.

preference towards $L_{\tau 3}$ conformations that allows for surface exposure of both, pyrimidine ring and cyano groups. Protonated TMC278 molecules assumed mostly the U conformations preferentially exposing the charged pyrimidine ring while burying the mostly hydrophobic wings.

Addition of co-solvent to the aggregate MD simulations

In addition to looking at a series of two-component systems discussed above, behavior of a three-component system where another co-solvent was included was also initiated. The third component added to the simulation was the universal co-solvent DMSO that was used in the biophysical studies discussed in the first chapter. Formation of drug aggregates in the absence of DMSO *in vitro* has not been successful (data not shown), for that reason, we believe that the additive is important for drug crystal dispersion and possibly water/aggregate interface formation. However, despite full dissolution of drug in 100% DMSO (or PEG400) formation of drug aggregates has been observed upon introduction of aqueous media. Also, the earlier presented biophysical studies had shown that the nature of the co-solvent does not change aggregate behavior under acidic conditions. To address this aspect of TMC278 aggregate formation, we designed a subset of MD simulations to determine whether the presence of DMSO alters behavior of drug molecules at the water/aggregate interface.

To study behavior of the TMC278/DMSO/water systems, a subset of control models was designed. Drug concentration in DMSO used in the *in vitro* experiments corresponds to approximately a 300 to one ratio of DMSO to one TMC278 molecule (20 mM drug concentration in 100% DMSO). Generation and monitoring of molecular dynamics for a system of such scale was not perceived as a feasible task. Therefore, we designed three MD models that could address simpler, more succinct questions such as whether DMSO molecules prefer “oil” or “water” or an interface between the two media in the context of our simulation. To evaluate behavior of a complex system including three components, at first a set of simpler control experiments was designed to test

behavior of DMSO: 1) DMSO in water; 2) TMC278 in DMSO. The three-component system included water and drug species in DMSO. Figure 45 summarizes the initial and the final MD simulation structures for two controls and one three component experimental system. The DMSO in water simulation result (Figure 45A) shows that

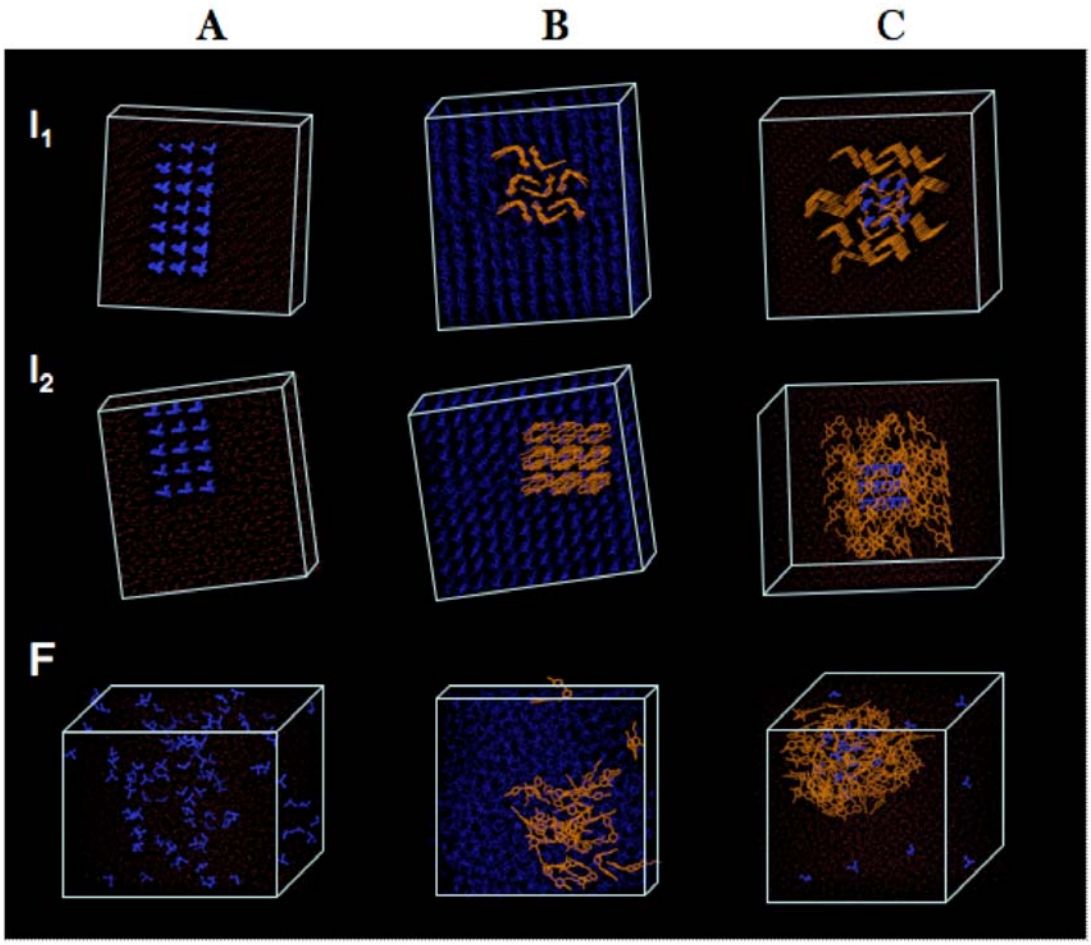


Figure 45. The starting and final structures of the three component aggregate MD simulations with DMSO. I_1 and I_2 correspond to two alternative views of initial MD model structure, F is a 7 ns equilibrated final structure. Color assignments: blue: DMSO; red: water oxygen; orange: TMC278. A. 96 DMSO molecules in a 50 x 50 x 50 Å water box. B. 24 DMSO molecules in a 50 x 50 x 50 Å DMSO box. C. 27 DMSO molecules in a TMC278 microcrystal (47 molecules) in a 60 x 60 x 60 Å water box.

DMSO prefers to be fully dispersed in aqueous solutions and this result agrees with *in vitro* behavior of DMSO in water, where final solutions are characteristically clear and

lack strong light scattering signal. The drug crystal in DMSO simulation (Figure 45B) has shown that over time the drug crystal is starting to fall apart. The rate of drug dissociation in DMSO is clearly greater than that in water, where the TMC278 crystal was still intact and most of the drug molecules were in the same position and orientation as the starting model after 7 ns of MD simulation. In the presently discussed simulation, after 7 ns the drug crystal is structurally rearranged and most of the molecules are solvated by DMSO. Based on the final MD structure it is apparent that if more simulation time were available, the drug crystal would have been fully dissolved in the DMSO box. Under similar experimental conditions *in vitro,* the drug concentration of the system is much lower, therefore it is plausible to speculate that if the size of the drug crystal was smaller or a longer simulation time was allowed, the drug crystal would fall apart completely. Results of the TMC278/DMSO simulation are consistent with *in vitro* results where dissolution of the dry TMC278 compound into a 100% DMSO at 50 mM drug concentration is not instantaneous, requires vortexing, yet yields a clear solution with low scattering signal in DLS experiments.

Results of the TMC278/DMSO/water MD simulation suggest that DMSO molecules mostly do not occupy the drug aggregate surface. As we know from the earlier described experiments on TMC278 in water, at 0% protonation TMC278 drug crystal is very stable in aqueous environment. Therefore in the present simulation, it is assumed that the oil/water interface coincides with the drug/water interface. The main focus of this simulation was to establish whether or not DMSO molecules also populate the drug/water interface. As shown in Figure 45C, DMSO molecules in the equilibrated MD simulation are either fully solvated or enclosed in the drug aggregate. MD trajectory

analysis has shown that DMSO molecules did not exhibit significant movement while being confined to the drug crystal. Once the drug molecule on the surface of the crystal moved enough to allow DMSO molecules to exit, the DMSO molecules initiated large movements and made contact with water. Movement of DMSO molecules directly into the aqueous phase instead of within the drug crystal suggests that packing induced interactions between drug molecules are stronger than DMSO driven crystal deformation seen in the control experiment. Also, these observations support our hypothesis that a solvent exposed DMSO molecule at the surface of an aggregate has tendency to move into the aqueous fraction instead of occupying the oil/water interface. Under *in vitro* conditions, we hypothesize that DMSO is present in the internal structure of the aggregate due to the large size of the system and lack of interaction with the outside aqueous environment. Therefore, based on the current MD simulation results we speculate that although the internal structure of the drug aggregate can contain DMSO, the drug-water interface most likely does not contain a significant fraction of DMSO molecules.

Due to the fact that the currently available three-component model is not representative of the *in vitro* conditions, additional experiments need to be conducted to verify earlier stated speculations regarding DMSO presentation at the surface of drug aggregates. Based on concentration information, we can design MD simulation in a step-by-step fashion. The first step of the MD simulation would mimic soaking a drug microcrystal in 100% DMSO solution (analogously to earlier shown control experiments). Post equilibration, a fraction of the system would need to be extracted to be soaked in a significantly larger water box. The number of molecules in both

simulations need to be kept at a minimum due to limitations associated with necessary equilibration time and corresponding CPU time requirements. Based on our current results, we hypothesis that the three-component system will display similar trends in DMSO behavior, therefore, showing movement of co-solvent located at the surface of an aggregate to the aqueous phase while those at the core would equilibrate within an aggregate. In order to answer questions of pH dependent behavior of an aggregate in the context of a three-component system, introduction of protonated species would also be beneficial.

Discussion

Aggregation of TMC278

Pharmacokinetic and biophysical studies of TMC278 suggest that one possible explanation for the exceptional oral bioavailability properties of the drug lies with its ability to form ~100-120 nm diameter, self-formulating NNRTI aggregates and their subsequent uptake into systemic circulation (Frenkel et al., 2005; Lewi et al., 2004). We hypothesized that formation of the monodispersed aggregate populations is a product of the intrinsic properties of the TMC278 molecule. We proposed that TMC278 aggregates, by size and morphology, resemble large micellar structures and/or microemulsions that are formed in the absence of conventional surfactants. Therefore, we hypothesized that the drug molecule is able to display surfactant-like properties under physiologically relevant conditions facilitating micelle or microemulsion-like formation without the aid of added surfactants. The computational study reported here support our hypothesis and suggest that under acidic conditions, protonation of TMC278 changes the compound's

behavior towards a more surface-active molecule with clear preference toward exposure of hydrophilic portions of the molecule and burial of the hydrophobic groups. The computer simulations also show a clear change in the aggregate size and shape and a transition from solid-like to liquid-like phase with increasing fraction of charged species in the simulations. The models also suggest that under solution conditions of the simulation, the aggregates form a microemulsion-like phase with a partially disordered internal structure (de Gennes and Taupin, 1982; Narang et al., 2007). The structural flexibility and the head group size of the surfactant molecules are known to be fundamentally important for micelle and interfacial surface formation (Ambade et al., 2007; Bhattacharya and Haldar, 2004; Haldar et al., 2005; Schlossman and Tikhonov, 2008). Smaller size, higher flexibility, and small electrostatic repulsions produce larger micelle structures and/or tighter surfactant molecule packing at the oil/water interface (Ambade et al., 2007; de Gennes and Taupin, 1982; Haldar et al., 2005; Narang et al., 2007). Just as in the study reported by Ambade et al, it has been observed that the flexibility of the surfactant molecule contributed to the final dimension of the aggregate with the most compact molecule forming a larger aggregated and/or micelle-like structure (Ambade et al., 2007) we believe that intrinsic flexibility of the TMC278 molecule, accompanied by its ability to assume amphiphilic properties under protonated conditions, similarly allows for more advantageous packing of drug molecules at the water/aggregate interface.

Aggregation at physiologically relevant conditions has been observed for many drugs and drug candidates, however, good bioavailability properties are not as frequently observed for these molecules. In the series of studies reported by Seidler, McGovern,

Feng, and other colleagues in the Schoichet group, formation of aggregate structures by drug and drug-like molecules has been correlated with promiscuous enzyme inhibition (Feng et al., 2005; Feng and Shoichet, 2006; McGovern et al., 2002; McGovern et al., 2003; Seidler et al., 2003) and false positive results in HTS screening for drug leads (Shoichet, 2004, 2006). In the recently published work by Coan et al. the drug aggregate composition is also compared to that of micelles (Coan and Shoichet, 2008). Using the drug concentration measurements and monomer-to-aggregate volume ratio calculations, the authors concluded that drug aggregates were densely packed. These findings support our hypothesis of micelle-like drug aggregate formation and justify the close packing of the drug aggregates observed in our simulations. Our results also suggest that mechanism of aggregate formation might be shared between promiscuous aggregators and molecules discussed in our work. While many drug molecules might form aggregate structures, we speculate that only those that are able to form these structures in a surfactant-independent fashion under physiologically relevant conditions might be able to follow absorption pathways associated with improved oral bioavailability. Conversely, we believe that drug or drug candidate molecules unable to display amphiphilic properties under physiologically relevant conditions, would tend to be less capable of forming stable, homogeneous populations of nanosize aggregates.

Conclusion

TMC278 is a BCS Class II compound with exceptional oral bioavailability properties that forms ~100 nm aggregates in surfactant-independent fashion at low pH and high drug concentration. We hypothesized that tendency of the drug molecule to form micelle-like structures under physiologically relevant conditions is attributed to its

ability to behave as a surfactant molecule at the aggregate/water interface. We conducted a series of molecular dynamic simulations of neutral and protonated TMC278 mixtures in water to test this hypothesis. Our simulation results supported our hypothesis and show that protonated drug molecules observe surfactant-like behavior at the aggregate/water interface by preferentially displaying hydrophilic parts of the molecule while burying hydrophobic ones. Charge density saturation, aggregate morphology, and drug aggregate packing analyses suggest that aggregate surface is populated by both neutral and protonated drug molecules. Neutral molecules exhibit structural variability that allows for better packing and liquid-like formation of the aggregate while protonated molecules affect aggregate shape and size. Based on our observations, we conclude that ability of the highly hydrophobic drug molecules to observe amphiphilic properties under conditions of low pH can improve their ability to form nanosize aggregate structures and therefore improve their oral bioavailability.

Methods

TMC278 and TMC125 infinite dilution MD simulations

Behavior of individual NNRTI molecules fully solvated in implicit water was evaluated using molecular dynamics. Two molecules, TMC278 and TMC125, under neutral and protonated conditions were subjected to temperature replica exchange molecular dynamics (REMD) methods (Felts et al., 2004; Gallicchio et al., 2008) using the computational package IMPACT with the OPLS-2005 force field (Banks et al., 1999; Kaminski et al., 2004; Kaminski et al., 2002; Maple et al., 2005; Stern et al., 1999; Stern et al., 2001). Starting molecular models for MD simulations were based on X-ray

structures of anhydrous TMC278, protonated TMC278, and TMC125 (reported in Chapter 2). Explicit solvent model used in MD calculations was SPC. Each system was minimized at 300K for 1 ps with a 0.001 ps time step using conjugated gradient. Ten nanosecond trajectories were collected with 0.001 ps time step. The walkers were set at the following temperatures: 300, 305, 310, 315, 320, 325, 330, 335, 340, 345, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, and 450 K using IMPACT (Banks et al., 2005). Trajectory information was recorded every 2 ps.

In order to study behavior of TMC278 and TMC25 in the MD simulations, observed conformations were categorized into four main groups based on rotations of four dihedral angles. Torsion angles, $\tau 1$ – $\tau 4$, were identified using previously established nomenclature (Figure 19). Drug conformations with $\tau 3$ and $\tau 4$ less than 90º were called "U" conformations (Figure 5a) and had been previously observed in crystallographic structures of the drug in protein complexes (wild-type and mutant)(Das et al., 2008), as well as in the small-molecule crystal structures (Frenkel et al., in preparation) and were referred to as either "horseshoe" or "butterfly." Drug conformations with $\tau 3$ greater and $\tau 4$ less than 90º resulted in rotation of wing I relative to pyrimidine ring and were called "$L_{\tau 3}$", while those with $\tau 3$ less and $\tau 4$ greater than 90º, where wing II was rotated relative to pyrimidine ring, were called "$L_{\tau 4}$" (Figure 40). The most torsionally flexed conformation of TMC278 was called "E," for extended, and was characterized by $\tau 3$ and $\tau 4$ being greater than 90º, where both wings were rotated relative to pyrimidine (Figure 40). The E conformation has been previously observed for a diaryltriazine analog of a DAPY NNRTI molecule, R120393, in complex with HIV-1 RT

and in a small-molecule crystal structure and has been called a “seahorse” conformation (Das et al., 2005; Frenkel et al., in preparation).

MD simulation of TMC278 aggregate surface

To study behavior of TMC278 in an aggregate five different starting MD models were designed. The five investigated models differ in the fraction of protonated TMC278 molecules that was varied from zero to ninety percent (see Table 11). The starting aggregate models were generated using small-molecule crystal structure coordinates of TMC278 crystal in anhydrous form (Frenkel et al., in preparation). One hundred TMC278 molecules were packed in the lattice according to the $P2_1/c$ space group symmetry of the crystals with final microcrystalline dimensions of approximately 35 x 35 x 45 Å. These models were inserted using the Maestro Program (Schrodinger, LLC) into pre-equilibrated simulation cells of 60 x 60 x 60 Å containing pre-equilibrated water system (SPC water model). Water molecules which overlapped with the solutes were removed. In the resulting model at least two layers of water molecules surrounding the generated drug crystals (see details in Table 11 and Figure 37). Depending on the model, a fraction of TMC278 molecules in the microcrystal were protonated at the N(2) atom of the molecule. The preferred protonation site at N(2) atom of the molecule was indicated by a recently obtained small-molecule crystal structure of protonated TMC278 (Frenkel et al., in preparation). Initial placement of protonated TMC278 molecules was performed in such a way to uniformly cover the surface of the microcrystal. If protonated drug molecules were present in the simulation, an equivalent number of chloride ions were included in the simulation to mirror observed solvent content in the small-molecule structure of the protonated TMC278 (Figure 37).

MD trajectories of TMC278 aggregate surface were obtained using the IMPACT molecular simulation program (Banks et al., 2005) with the 2005 implementation of the OPLS-2005 force field (Banks et al., 1999; Kaminski et al., 2004; Kaminski et al., 2002; Maple et al., 2005; Stern et al., 1999; Stern et al., 2001) (Schrodinger, LLC) under periodic boundary conditions and constant temperature and pressure at 298.15 K and 1 atm (Banks et al., 2005; Jorgensen et al., 1996). All covalent bonds were constrained using the SHAKE algorithm, with a relative tolerance of 10^{-7}. The MD time-step was set to 2 fs. The 60 and 90% protonated models were simulated for 7 ns and the 0, 10, and 30% protonated models for 9 ns. System coordinates were saved every 20 ps after a 2 ns initial system equilibration period. The resulting trajectories were analyzed using the Maestro program and associated utilities.

Aggregate cluster analysis

A molecule was considered part of the aggregate if it made significantly fewer water contacts than fully solvated molecules. To this end we computed for each molecule *m* in each MD snapshot a solvation parameter S_m defined as the number of water molecules within 4 Å radius of any atom of the molecule. We defined a fully buried molecule in the core of the aggregate as having $S_m = 0$. S_m values greater than 20 correspond to fully dissociated drug molecules; $9 < S_m < 20$ correspond to molecules on the surface of an aggregate, and $0 < S_m < 10$ correspond to molecules that compose the interface between the surface layer and the core of an aggregate (Table 1). The aggregate size is expressed in terms of average number of molecules in an aggregate (N_{agg}), and average number of surface, core, and charged molecules as reported in Table 11.

Aggregate surface characterization

MD trajectories were evaluated in terms of aggregate morphology, TMC278 torsional flexibility, aggregate packing, and drug molecule surface contribution as a function of drug location within an aggregate. Evaluation of aggregate morphology was done by visual inspection using the Maestro graphical interface in combination with solvent accessible surface area (SASA) estimation based on the periodic boundary-related contacts between models, if such were present, using a 1.4 Å radius probe.

The packing of the molecules in an aggregate was evaluated using ring-to-ring distance distribution analysis for 0, 10, and 30% models based on the distances between the centers of mass for selected ring pairs, sampling distances within a 10 Å radius. Both intra- and intermolecular ring-to-ring distances were calculated. Based on stacking interactions of the crystal packing observed in the TMC278 anhydrous crystalline structure, we selected ring pairs for monitoring disruption of crystal packing as a function of protonation (Figure 39). Distances of 3.5 ± 0.5 Å correspond to intermolecular stacking interactions between wing I and wing II; and the symmetry-related wing II rings in the TMC278 crystal packing.

The contribution of each atom to the aggregate solvent accessible surface area (SASA) was computed with SURFV software using a 1.4 Å rolling probe (Nicholls et al., 1991; Sridharan et al., 1992). Per atom surface contributions were normalized against the maximum SASA (mSASA) of an atom based on its most solvent exposed TMC278 conformation. Average SASA and number of surface charged molecules were used to estimate average aggregate surface charge density, reported in Table 11. Figure 42

pertain results of the hydration analysis, where the color gradient ranging from dark blue corresponds to small percent contributions to the surface and red (highest solvation) is used to accent the least and the most hydrated atoms of the molecules contributing to aggregate surface.

DMSO containing drug aggregate surface MD simulation models

Three-component MD simulations containing water, drug, and DMSO molecules were designed to evaluate behavior of DMSO at the water/aggregate interface. Under experimental conditions *in vitro*, DMSO was used as a solvent for TMC278; therefore, the solvent model for DMSO was generated using density and molecular weight information. It was determined that a 5 × 5 × 5 Å cell fits 0.96 molecules of DMSO, using this information, a 60 × 60 × 60 Å lattice was constructed using P1 symmetry followed by 1000 steps of minimization at 298K. When minimized DMSO solvent model was used, any DMSO molecule overlapping any atom of the solute molecule was removed.

Initially, the two-component control experiments, testing behavior of DMSO either in water or in the presence of the dug molecules only, were evaluated. The DMSO in water control included a block of 96 DMSO molecules soaked in a 50 × 50 × 50 Å equilibrated SPC water box (Figure 45). The second control testing behavior of drug molecules in DMSO was created by soaking a TMC278 anhydrous crystal (27 molecules) into a 50 × 50 × 50 Å DMSO box (Figure 45). The three-component DMSO/drug/water experimental system was composed of 47 DMSO molecules embedded in a 27 molecule

TMC278 anhydrous crystal, soaked in the equilibrated SPC water box (60 × 60 × 60 Å water box). Design of the-three component model was not meant to reflect our perception of the biological system but to allow exploration of DMSO behavior at the water/aggregate interface. All three models were subject to 7 ns of MD simulation using Impact using OPLS-2005 force field at 298 K.

Chapter 4. MD simulation of NNRTI binding pocket in complex with TMC278

Synopsis

Previously published crystallographic studies of TMC278 in complex with HIV-1 RT had revealed that the drug molecule and related DAPY NNRTIs display multiple binding modes in the NNIBP of the enzyme (Das et al., 2008; Das et al., 2005; Lewi et al., 2003). Though multiple, static 'snapshots' of alternative conformations within the NNIBP had been observed, the ligand dynamics in the pocket are still unclear. TMC278 is one of the most potent and resistance-resilient inhibitors of the wild-type and multiple mutant forms of HIV-1. Therefore, to study behavior of TMC278 in the NNIBP, a number of MD simulations had been designed. 4 ns MD trajectories of TMC278 in complex with the wild-type and mutant (Y181/K103N) HIV-1 RT in implicit solvent were assessed in terms of rotational and translational movements of the drug. MD trajectories were evaluated using conformational sampling and principal component analysis (PCA). Experimental results support strategic flexibility hypothesis and suggest that TMC278 samples greater rotational freedom ('wiggling') in the mutant-binding pocket than in the wild type. Translational movements ('jiggling') were observed at a lesser extent and significant displacements were speculated to require a longer simulation time.

Introduction

A centrally important property of TMC278 is its ability to inhibit a broad range of drug-resistant HIV-1 RT mutant strains. Previous crystallographic studies of TMC278 and related NNRTIs in complex with wild-type and mutant HIV-1 RT had shown that the compounds could bind with significant conformational variation in the binding pocket of the enzyme (Das et al., 2008; Das et al., 2004; Das et al., 2005). It has been hypothesized that by means of conformational and positional (rotational and translational) changes ('wiggling' and 'jiggling'), the TMC278 molecule is able to accommodate conformational changes in binding pocket landscape of the drug-resistant mutant HIV-1 RTs. The small molecule X-ray structure analysis presented in the second chapter also provides evidence of broad conformational sampling by the drug molecules in both free (infinite dilution) and constrained (crystal packing and aggregate core) environments (Frenkel et al., in preparation; Frenkel et al., submitted). Though multiple static conformations had been detected, the dynamic information on the movement of TMC278 molecule in the receptor-bound state has not been yet explored in detail.

Ligand flexibility in the binding pocket had been viewed as both a benefit and a detriment to the substrate/enzyme binding interaction. On one hand, ligand flexibility can worsen the substrate/enzyme energetics by increasing the ligand entropy; on the other, it can enhance the drug binding by allowing for better accommodation of the binding surface. The role of strategic flexibility in the substrate-enzyme interaction has presented from both perspectives in different substrate/enzyme systems. Examples of such systems include HIV-1 protease (Hornak and Simmerling, 2007) and S-adenosyl-L-homocysteine hydrolase (SAHase) (Polak et al., 2004), to name a few. Flexibility of the

HIV-1 protease active site region is believed to be a product of conformational sampling explored by two β-hairpins, or "flaps" comprising the binding pocket (Lapatto et al., 1989a; Lapatto et al., 1989b; Navia et al., 1989). The structural flexibility of the flaps is attributed to the large glycine content, especially in the mutants (Perryman et al., 2006). One of the models suggests that the structural flexibility of the pocket in the mutants limits ability of the ligands to access the binding pocket by preventing favorable interactions between the drug and surrounding residues for proper enzymatic mechanics (Martin et al., 2005) while another group proposes that enhanced flexibility of the mutant structures results in drug resistance due to inability of the less flexible synthetic inhibitors to adapt their conformations to that of the mutant protease binding pocket landscape (Luque et al., 1998).

Examples of recently introduced inhibitors with improved activity attributed to increased ligand flexibility are nucleoside analogs called fleximers (Polak et al., 2004; Seley et al., 2003). The nucleoside fleximers are composed of a purine nucleobase split into two components, imidazole and pyrimidine, designed to inhibit activity of SAHase (Polak et al., 2004; Seley et al., 2003). It has been shown that the nucleoside fleximer with increased flexibility was able to observe a larger number of conformations that reflected the conformational changes experienced by the enzyme (Seley et al., 2003).

Interestingly, both flexibility models had been claimed to play a role in mutation mechanisms, either as a mechanism for development of drug resistance by the enzyme or as a mutation resistant mechanism employed by the substrate. The main concept that brings the two flexibility models together is that an enzyme and a substrate are pieces of an interactive molecular puzzle in which strategic flexibility plays a critical role.

Previously described studies on protein and ligand flexibility suggest that both partners can be highly flexible and dynamic, therefore, when flexibility of the therapeutic target is reflected in the flexibility of the therapeutic agent, a more potent and a more resilient drug candidate is developed (Seley 2005). In the present work, molecular mechanics are employed to look at the dynamics of the bound TMC278 in the pocket of both wild-type and mutant HIV-1 RT with the goal of defining the extent of flexibility exhibited by both 'pieces of the molecular puzzle.'

Results and Discussion

'Wiggling' and 'jiggling' as described by Das et al., correspond to ability of the drug molecule to sample multiple conformations and positions in the NNIBP of the enzyme. Though multiple binding modes had been observed, dynamics of the bound TMC278 in the pocket of HIV-1 RT had not been yet explored in detail. The study presented here focused on the binding dynamics of the drug and NNIBP residues of HIV-1 RT using computational methods. As rotational and translational changes of the ligand were claimed to play part in mutation resistance, computational study included analysis on both wild-type and Y181C/K103N mutant HIV-1 RTs.

Structural variability of TMC278 in the context of an unconstrained (infinite dilution and water/aggregate interface) and a constrained (small molecule crystal packing and HIV-1 RT complex) molecular environment had been studied previously by monitoring of the torsional changes in the drug molecule (Das et al., 2008; Frenkel et al., in preparation; Frenkel et al., submitted). Similarly, conformational changes attributed to rotations at dihedral angles $\tau 1 - \tau 5$ were monitored in TMC278/RT MD simulations (Figure 46). The torsional sampling probability plots reported in Figure 46 suggest that

in the TMC278/RT MD trajectories dihedral rotations observed for drug molecules at $\tau1 - \tau4$ were mostly conserved in both wild-type and mutant complexes. At all four torsion angles, the deviation from the median was averaging around ± 30°. The same did not hold true for the sampling of the dihedral angle $\tau5$ where in the wild-type system 100% of the molecules were at 0°, while in the mutant only 60% of the conformations contributed to this group. The other 40% of the structures in the mutant simulation sampled $\tau5$ angle conformations at ~160°. We investigated the energetic differences for TMC278 structures in folded conformation with different cyanovinyl orientations by varying $\tau5$ from 0° to 360° and measuring their corresponding potential energies. As shown in Figure 47, TMC278 in "horseshoe" conformation with $\tau5$ at 0° has only a slightly lower energy than that at 160° (<0.5 kcal/mol), therefore suggesting that both conformational pools are energetically favorable. Though the energy required to rotate $\tau5$ from 0° to 150-160° had not been evaluated in this study, based on the frequent and random sampling of these conformation in the mutant MD trajectory, it is hypothesized that the energy barrier was low.

'Jiggling' or repositioning of the drug molecule in the pocket was studied using principal component analysis (PCA). PCA is a mathematical method used to process a large amount of data to identify possible correlations. By means of PCA the size of the final dataset is greatly reduced and the most significant movements are identified. The output of the PCA analysis is a collection of frames that are a set of average structures each corresponding to a point on the principal component axis. The first principal component axis reports structures with highest degree of variance and thus of most significance. To evaluate PCA output using experimental structures, three points on the

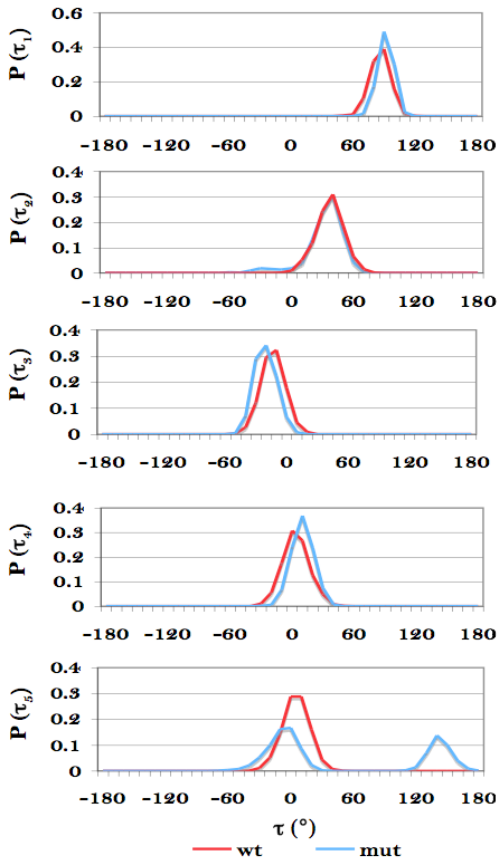


Figure 46. The probability distribution plots for torsional angles τ1 – τ5 of TMC278 in TMC278/RT MD simulations. The color assignment: red: wild type; blue: mutant Y181C/K103N RT.

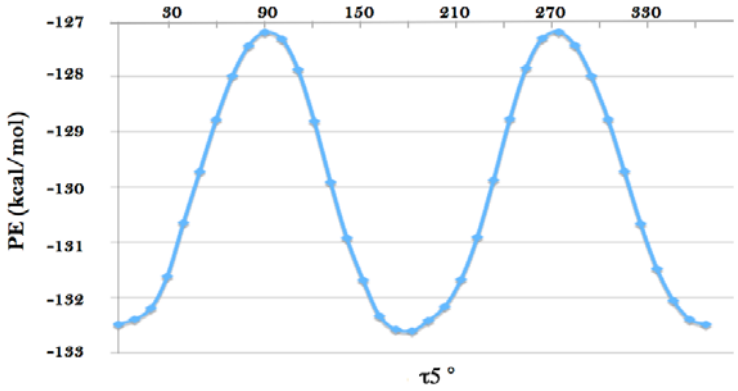


Figure 47. The plot of potential energy distribution for the TMC278 dihedral angle, τ5. Potential energy calculations were performed using the MacroModel coordinate scan utility (see Methods for details).

collection of MD trajectory structures that correspond to the selected projections on the first PCA axis were chosen: the origin, the negative extreme, and the positive extremes (Figure 48). The top ten MD structures with RMSD values under 0.25 Å for the TMC278 coordinates were selected and used for analysis of drug movement in the pocket. The first principal component axis allowed characterization and comparison of motion displayed by TMC278 in the two different NNIBPs of RT within the timeframe of the MD experiments. As can be seen from the superposition of structures corresponding to the three points on the first PCA axis shown in Figure 48, TMC278 in the wild-type pocket

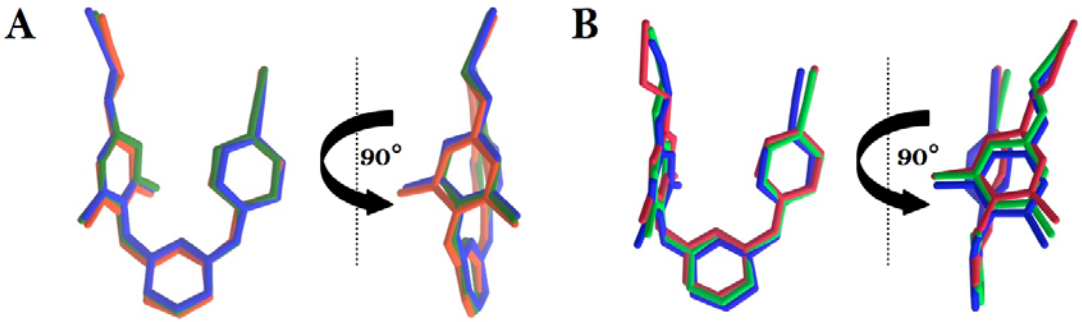


Figure 48. First PCA axis projections for the wild-type and the mutant TMC278/RT MD simulations. **A.** Structures of the origin, the positive, and the negative extremes of the PCA analysis for the wild-type TMC278/RT MD simulation. **B.** Structures of the origin, the positive, and the negative extremes of the PCA analysis for the mutant TMC278/RT MD simulation. Color assignment: green: origin; red: positive extreme; blue: negative extreme.

of RT did not show significant change in either its position (~0.1 Å displacement between the atoms of pyrimidine ring in the extremes and the origin) or in the predominant conformation (Figure 48A). In the mutant MD simulation, just as it was discussed earlier, significant changes were associated mostly with the rotation of the cyanovinyl moiety as well as a ~0.3 Å displacement from the

origin (Figure 48B). In addition to displacement from the first PCA origin, displacement from the crystallographically obtained TMC278 coordinates in the pocket was also calculated. The average displacement of the positive and negative PCA extremes from the crystallographic structures was determined to be 0.2 Å in the wild type and 0.5 Å in the mutant. Though translational movements observed in the MD simulations were small between the two compared systems, TMC278 in complex with the mutant RT sampled more pronounced 'jiggling' than its counterpart.

In the molecular puzzle model, both molecular pieces (a substrate and an enzyme) move to maximize their favorable interactions. To assess extent of interaction between the drug molecule and the surrounding residues of the NNIBP, the buried surface analysis on the three points of the first PCA axis was initiated (Figure 49). As it was determined earlier in the wild-type TMC278/RT simulation TMC278 did not change significantly either its conformation or its position and as can be seen from the buried surface plots shown in Figure 49, the nature of interactions between the drug and the protein did not change significantly as a result of these movements either. The largest buried surface in these structures is the one that corresponding to the location of the hydrogen bond between the linker II nitrogen and the carbonyl oxygen of K101. Inhibitor-protein interactions were also seen between wing II and V106, linker I region and the main chain atoms of Y181, and wing II and Y318. In contrast to the wild-type simulations, the three points on the first PCA axis in the mutant simulation

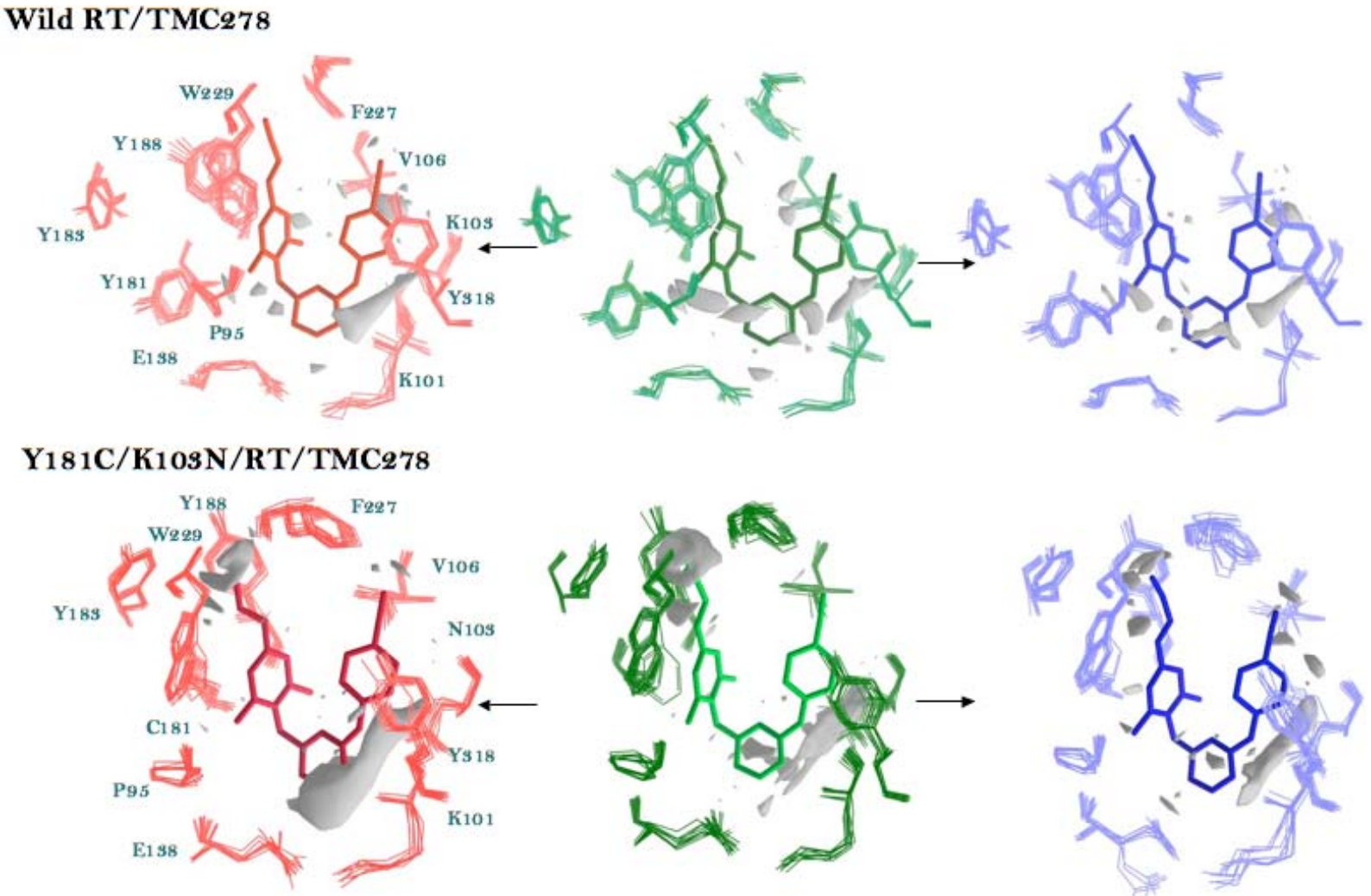


Figure 49. Buried surface area comparison for the three points on the first PCA axis in the wild-type and mutant RT/TMC278 MD simulations. Color assignment: green – PC origin, blue – PC negative extreme, red – PC positive extreme.

display significant differences between each other. As with the wild-type, the largest buried surface area corresponds to the hydrogen bond between the linker II position and carbonyl of K101. Yet in contrast to the wild-type, the mutant structures also show a substantial change in the interactions between the cyanovinyl moiety of the drug and the residues surrounding it. As can be seen from the figure, F227, Y188, and W229 surround the cyanovinyl group creating a hydrophobic tunnel as described in the crystal structure analysis (Das et al., 2008). At the first PCA origin, the position of the cyano group is such that allows for the most extensive interaction with the surrounding residues. In the positive extreme the cyanovinyl rotated and made contacts with W229 and Y188, whereas interactions with F227 were significantly reduced. In attempt to preserve a favorable interaction, the F227 side-chain phenyl ring compensated for rotation of the cyanovinyl group by moving towards it ~0.5 Å. In the negative direction, cyanovinyl moves down by 1.5 Å and therefore also diminishes its interaction with surrounding residues. To compensate, F227, Y183, and Y188 move ~0.4 Å towards it. In summary, PCA findings suggest that the hydrogen bond at the linker II position ‘tethers’ the drug in the pocket in the most conservative fashion, which explains the preserved buried surface interaction in all PCA snapshots. At the same time, due to changes in the pocket and ability of the molecule to alter its conformation via torsional changes, the wing I moiety is forced to seek alternative orientations, which in this simulation results in a sliding motion up and down the hydrophobic tunnel.

The PCA results suggest that the movement of the surrounding NNIBP residues sometimes compensates for the movement of the drug molecule. To account for the breadth of the non-compensatory movement in the TMC278/RT MD simulations, the intermolecular distances between ligand and protein atoms were measured throughout the MD trajectories. As depicted in Figure 50, selection of the pair-wise distances was conducted so that all parts of the drug molecule were monitored by multiple pair-wise interactions. For example, the movement of pyrimidine ring was evaluated in terms of distances between pyrimidine atoms and three binding pocket residues: L100, V179, and K101 (Table 12). Generally protein atoms used in the pair-wise distance measurements were either highly conserved throughout the MD trajectory (with RMSD <0.2 Å) or were closely packed and therefore expected to form extensive interaction. In both cases, the difference in interatomic distances and maximum displacement values were expected to reflect an extent of uncoordinated movement

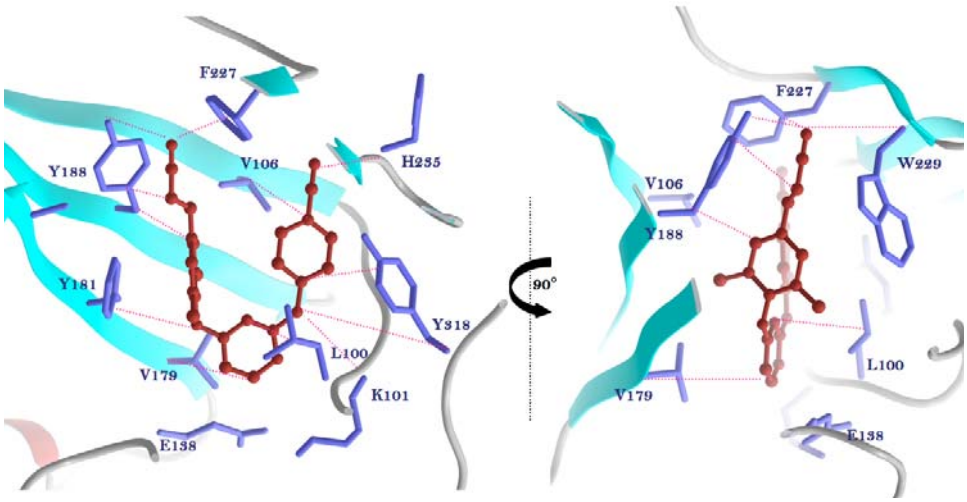


Figure 50. Interatomic distances used to evaluate non-compensatory movement of TMC278 in the wild type HIV-1 RT binding pocket. Atom designation used in the pair-wise distance assignment is listed in the Methods section.

Interatomic distances (d)	wtRT/TMC278 d ± Δ (Å)	K103N/Y181C/RT/TMC278 d ± Δ (Å)
Y/C181 - Linker I (N4)	4.7 ± 0.5	5.0 ± 0.6
L100 - Pyrimidine	4.4 ± 0.7	4.6 ± 0.7
V179 - Pyrimidine	5.7 ± 0.8	5.6 ± 0.7
K101 - Pyrimidine	3.1 ± 0.5	3.1 ± 0.6
K101 - Linker II (H bond)	2.8 ± 0.4	2.8 ± 0.4
Y318 - Linker II	6.5 ± 0.6	7.5 ± 0.7
Y318 - Wing II	3.6 ± 0.5	4.4 ± 1.1
V106 - Wing II	5.8 ± 0.6	5.1 ± 0.8
H235 - Cyano Wing II	3.9 ± 0.9	3.7 ± 0.7
F227 - Cyano Wing I	3.7 ± 0.6	4.2 ± 0.9
Y188 - Cyano Wing I	5.0 ± 1.0	4.1 ± 0.9
W229 - Cyano Wing I	4.7 ± 1.1	4.3 ± 0.7
Y188 - Wing I	3.4 ± 0.4	3.7 ± 0.8

Table 12. The list of pair-wise interatomic distances between the NNIBP residues and TMC278 in the TMC278/RT MD simulations. Atom assignment used to determine average distance (d) and maximum displacement (Δ) are reported in the Methods section and illustrated in Figure 50.

between selected atoms. As can be seen from Table 12, uncoordinated movements were widely sampled in both MD simulations. In general, the wild-type displacement differences were slightly smaller (average Δ ~0.7 Å) than those observed in the mutant complex (average Δ ~0.8 Å), with exception for interaction between W229 and the cyano group on Wing I where in the wild-type complex the maximum displacement was 1.1 Å and in the mutant simulation 0.7 Å. These observations agree with the first PCA results, where the movement of the cyano group in the mutant was coordinated with W229 whereas in the wild-type complex the two moved more independently.

Rigid body translational movements of the drug in the context of the TMC278/RT MD simulation presented here were smaller than positional differences observed between

crystallographic structures of different TMC278/RT complexes. One possible explanation of this result is that the length of the MD simulations in the present study was not sufficient to observe full dynamic range of motion sampled by the drug post binding. To evaluate behavior of TMC278 in the binding pocket of RT exhaustively, a longer trajectory using of REMD method is recommended.

Conclusion

In summary, conformational sampling attributed to changes at dihedral angles had been observed in both wild-type and mutant TMC278/RT complexes. Changes at $\tau 1 - \tau 4$ were mostly conserved between two systems, whereas rotation of the cyanovinyl group by $\tau 5$ had shown larger conformational sampling in the mutant structure. By means of PC analysis, the most significant movements in two simulations were analyzed. In both cases translational changes were determined to be small, suggesting that either the movement post binding is small or that the length of the MD simulation has to be longer.

Methods

MD simulation of TMC278 in complex with HIV-1 RT

To study conformational sampling by TMC278 while in complex with HIV-1 RT, X-ray structures of wild type (PDB accession code 2ZD1) and K103N/Y188C-RT mutant (3BGR) without solvent molecules were used as starting models for MD simulation. Starting models were subjected to 1000 steps of minimization using the OPLS-2005 force

field in IMPACT at 298 K. For purposes of CPU time conservation, only the binding pocket of the HIV-1 RT that included any residue within a 5 Å radius from any atom of the ligand, was subjected to unconstrained MD calculation. To accommodate local residue movement and to prevent unfavorable clashes, a constrained buffer zone, defined as all residues within 5 Å radius from any atom of the binding pocket, was also included in the calculations. The rest of the atoms in protein remained frozen and therefore were not part of MD calculations.

MD trajectories were generated for the wild and the mutant HIV-1 RT in complex with TMC278 using the OPLS-2005 force field (Banks et al., 1999; Kaminski et al., 2004; Kaminski et al., 2002; Maple et al., 2005; Stern et al., 1999; Stern et al., 2001) (Schrodinger, LLC) and the implicit water model AGB in IMPACT under conditions of constant temperature and pressure. Temperature and pressure were kept at 298 K and 1 atm respectively. All covalent bonds were constrained using the SHAKE algorithm, with a relative tolerance of 10^{-7}. The MD time-step was set to 1 fs. Trajectory statistics and coordinates were recorded every 1000 steps after the first 2 ns of initial equilibration period. 2000 structures were included in data analysis using the Maestro program and associated utilities.

MD trajectory analysis

Similarly as for the conformational analysis reported in Chapter 3 for infinite dilution and aggregate MD simulations, the sampling of conformational space by TMC278 in the bound state was investigated using the monitoring of dihedral angles $\tau 1 - \tau 5$. The convention for dihedral angle assignments was stated earlier in Chapter 2, Figure 32. Along with monitoring of the dihedral angle sampling, potential energy

calculations were performed to evaluate effect of conformational sampling by τ5 in the wild-type complex. The coordinate scan utility in the MacroModel program from the Maestro software package (Schrodinger, LLC) was utilized to calculate potential energy of TMC278 conformations with alternative τ5 conformations. τ5 was rotated from 0° to 360° in 5° increments whereas all other angles in the drug structures were fully constrained.

In addition to looking at torsion angle changes sampled by the drug molecule in the binding pocket of the protein, the changes in positioning of the ligand in the pocket were also assessed. Non-compensated translational movements of the ligand in the binding pocket were investigated using interatomic pair-wise distances between selected protein atoms and those of TMC278. Protein atoms were included in the pair-wise distance analysis if they displayed either a relatively small range of motion throughout the MD trajectory or were closely associated with any heavy atom of the drug molecule. The following interatomic distances between protein and TMC278 atoms were used to evaluate MD trajectories for both wild-type and mutant MD simulations: Y/C181 (CA) – Linker I (N4), L100 (CG) – Pyrimidine (N1), V179 (CA) – Pyrimidine (C3), K101 (NH) – Pyrimidine (N2), K101 (CA) – Linker II (N3), Y318 (CA) – Linker II (N3), Y318 (CZ) – Wing II (C5), V106 (CB) – Wing II (C8), H235 (CA) – Cyano Wing II (N5), F227 (CG) – Cyano Wing I (N6), Y188 (OH) – Cyano Wing I (N6), W229 (CA) –Cyano Wing I (N6), Y188 (CG) – Cyanovinyl (C20), Y188 (CB) – Wing I (C14) (Figure 50). Pair wise distances were reported as average distances (d) with maximum displacement (Δ) values in Table 12. Δ is the difference between the largest or the smallest distance value

in the trajectory for a given pair of atoms and the average interatomic distance for a given pair wise interaction.

To select for structures that reflect the most coordinated movement of the drug within the binding pocket of the enzyme and to decrease the number of structures included in the analysis, a principal component analysis (PCA) was employed (Garcia, 1992) and conducted with assistance from Tom Eck. For PCA analysis, Cartesian coordinates of TMC278 heavy atoms were considered. Full 2 ns trajectory was evaluated. PCA of MD trajectory was performed using the open source statistical analysis package R V2.8.1 (Ripley, 2001, Team, 2006 #174), and PC results were extracted using the R utility, Bio3D (Grant et al., 2006) and custom codes to process the generated snapshots was written by Tom Eck. PCA results were visualized using the open source visualization software VMD (Humphrey et al., 1996). Three points on the first PC axis (origin, positive, and negative extremes) were used as the reference structures to search the wild-type and mutant complex MD trajectories for the most representative experimental coordinates. The top ten MD trajectory frames with TMC278 structures within 0.25 Å RMSD of the PCA structures were selected for each of the three first PC axis reference points and used in the subsequent analysis.

To study packing motifs of TMC278 in the binding pocket, the buried surface between binding pocket residues and the drug molecule was calculated. Due to the large size of the MD simulation dataset, surface calculations on the full trajectories were deemed to be unreasonable. Therefore structures corresponding to the three points on the first PCA axis were used instead. The buried surfaces were calculated between pocket

residues and TMC278 atoms. Visualization of the buried surfaces was achieved by using the intersection Boolean operation in the Maestro software package (Schrodinger, LLC).

References

Ahr, G., Voith, B., and Kuhlmann, J. (2000). Guidances related to bioavailability and bioequivalence: European industry perspective. Eur J Drug Metab Pharmacokinet *25*, 25-27.

Allen, F.H., Bird, C.M., Rowland, R.S., Harris, S.E., and Schwalbe, C.H. (1995). Correlation of the hydrogen-bond acceptor properties of nitrogen with the geometry of the Nsp2-Nsp3 transition in the R1(X-)C-NR2R3 substructures: reaction pathway for the protonation of nitrogen. Acta Cryst *B51*, 1068-1081.

Ambade, A.V., Aathimanikandan, S.V., van der Poll, D., and Thayumanavan, S. (2007). Smaller building blocks form larger assemblies: aggregation behavior of biaryl-based dendritic facial amphiphiles. The Journal of organic chemistry *72*, 8167-8174.

Amidon, G.L., Lennernas, H., Shah, V.P., and Crison, J.R. (1995). A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. Pharm Res *12*, 413-420.

Ansede, J.H., and Thakker, D.R. (2004). High-throughput screening for stability and inhibitory activity of compounds toward cytochrome P450-mediated metabolism. J Pharm Sci *93*, 239-255.

Arts, E.J., Miller, J.T., Ehresmann, B., and Le Grice, S.F. (1998). Mutating a region of HIV-1 reverse transcriptase implicated in tRNA(Lys-3) binding and the consequences for (-)-strand DNA synthesis. The Journal of biological chemistry *273*, 14523-14532.

Banks, J.L., Beard, H.S., Cao, Y., Cho, A.E., Damm, W., Farid, R., Felts, A.K., Halgren, T.A., Mainz, D.T., Maple, J.R., *et al.* (2005). Integrated Modeling Program, Applied Chemical Theory (IMPACT). Journal of computational chemistry *26*, 1752-1780.

Banks, J.L., Kaminski, G.A., Zhou, R., Mainz, D.T., Berne, B.J., and Friesner, R.A. (1999). Parametrizing a polarizable force field from ab initio data. I. The fluctuating point charge model. J Chem Phys *110*, 741-754.

Barat, C., Lullien, V., Schatz, O., Keith, G., Nugeyre, M.T., Gruninger-Leitch, F., Barre-Sinoussi, F., LeGrice, S.F., and Darlix, J.L. (1989). HIV-1 reverse transcriptase specifically interacts with the anticodon domain of its cognate primer tRNA. The EMBO journal *8*, 3279-3285.

Bhattacharya, S., and Haldar, J. (2004). Thermodynamics of micellization of multiheaded single-chain cationic surfactants. Langmuir *20*, 7940-7947.

Bohm, H.J., Klebe, G., Brode, S., and Hesse, U. (1996). Oxygen and nitrogen in competitive situations: which is the hydrogen-bond acceptor? Chem Eur J *2*, 1509-1513.

Boyer, J.C., Bebenek, K., and Kunkel, T.A. (1992). Unequal human immunodeficiency virus type 1 reverse transcriptase error rates with RNA and DNA templates. Proc Natl Acad Sci U S A *89*, 6919-6923.

Brameld, K.A., Kuhn, B., Reuter, D.C., and Stahl, M. (2008). Small molecule conformational preferences derived from crystal structure data. A medicinal chemistry focused analysis. Journal of chemical information and modeling *48*, 1-24.

Caldwell, G.W., Easlick, S.M., Gunnet, J., Masucci, J.A., and Demarest, K. (1998). In vitro permeability of eight beta-blockers through Caco-2 monolayers utilizing liquid chromatography/electrospray ionization mass spectrometry. J Mass Spectrom *33*, 607-614.

Camilleri, M., Colemont, L.J., Phillips, S.F., Brown, M.L., Thomforde, G.M., Chapman, N., and Zinsmeister, A.R. (1989). Human gastric emptying and colonic filling of solids characterized by a new method. Am J Physiol *257*, G284-290.

Charman, W.N.A., and Stella, V.J. (1986). Estimating the maximal potential for intestinal lymphatic transport of lipophilic drug molecules. Int J Pharm *34*, 175-178.

Charman, W.N.A., Stella, V.J., and Eds. (1992). Lymphatic transport of drugs (CRC Press).

Chen, H., and Langer, R. (1998). Oral particulate delivery: status and future trends. Adv Drug Deliv Rev *34*, 339-350.

Chen, L. (1995). Area dependence of the surface tension of a Lennard-Jones fluid from molecular dynamics simulations. J Chem Phys *103*, 10214-10216.

Chen, Z., Cheng, X., Haishi, C., Cheng, P., and Wang, H. (2007). Dissipative particle dynamics simulation of the phase behavior and microstructure of CTAB/octane/1-butanol/water microemulsion. Colloid Surf A: Physicochem Eng Aspects *301*, 437-443.

Coan, K.E.D., and Shoichet, B.K. (2008). Stoichiometry and physical chemistry of promiscuous aggregate-based inhibitors. J Am Chem Soc *130*, 9606-9612.

Coffin, J.M. (1995). HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science *267*, 483-489.

Das, K., Bauman, J.D., Clark, A.D., Jr., Frenkel, Y.V., Lewi, P.J., Shatkin, A.J., Hughes, S.H., and Arnold, E. (2008). High-resolution structures of HIV-1 reverse transcriptase/TMC278 complexes: strategic flexibility explains potency against resistance mutations. Proc Natl Acad Sci U S A *105*, 1466-1471.

Das, K., Clark, A.D., Jr., Lewi, P.J., Heeres, J., de Jonge, M.R., Koymans, L.M., Vinkers, H.M., Daeyaert, F., Ludovici, D.W., Kukla, M.J., *et al.* (2004). Roles of conformational and positional adaptability in structure-based design of TMC125-R165335 (etravirine) and related non-nucleoside reverse transcriptase inhibitors that are highly potent and effective against wild-type and drug-resistant HIV-1 variants. J Med Chem *47*, 2550-2560.

Das, K., Ding, J., Hsiou, Y., Clark, A.D., Jr., Moereels, H., Koymans, L., Andries, K., Pauwels, R., Janssen, P.A., Boyer, P.L., *et al.* (1996). Crystal structures of 8-Cl and 9-Cl TIBO complexed with wild-type HIV-1 RT and 8-Cl TIBO complexed with the Tyr181Cys HIV-1 RT drug-resistant mutant. Journal of molecular biology *264*, 1085-1100.
Das, K., Lewi, P.J., Hughes, S.H., and Arnold, E. (2005). Crystallography and the design of anti-AIDS drugs: conformational flexibility and positional adaptability are important in the design of non-nucleoside HIV-1 reverse transcriptase inhibitors. Prog Bioph and Mol Bio *88*, 209-231.

Das, K., Sarafianos, S.G., Clark, A.D., Jr., Boyer, P.L., Hughes, S.H., and Arnold, E. (2007). Crystal structures of clinically relevant Lys103Asn/Tyr181Cys double mutant HIV-1 reverse transcriptase in complexes with ATP and non-nucleoside inhibitor HBY 097. J Mol Bio *365*, 77-89.

De Clercq, E. (2004). Non-nucleoside reverse transcriptase inhibitors (NNRTIs): past, present, and future. Chemistry & biodiversity *1*, 44-64.

de Gennes, P.G., and Taupin, C. (1982). Microemulsions and the Flexibility of Oil/Water Interfaces. J Phys Chem *86*, 2294-2304.

Derecskei, B., Derecskei-Kovacs, A., and Schelly, Z.A. (1999). Atomic-level molecular modeling of AOT reverse micelles. 1. The AOT molecule in water and carbon tetrachloride. Langmuir *15*, 1981-1992.

Desai, M.P., Labhasetwar, V., Amidon, G.L., and Levy, R.J. (1996). Gastrointestinal uptake of biodegradable microparticles: effect of particle size. Pharm Res *13*, 1838-1845.

Ding, J., Das, K., Moereels, H., Koymans, L., Andries, K., Janssen, P.A., Hughes, S.H., and Arnold, E. (1995). Structure of HIV-1 RT/TIBO R 86183 complex reveals similarity in the binding of diverse nonnucleoside inhibitors. Nat Struct Biol *2*, 407-415.

Dressman, J.B., and Reppas, C. (2000). In vitro-in vivo correlations for lipophilic, poorly water-soluble drugs. Eur J Pharm Sci *11 Suppl 2*, S73-80.

Eicke, H.-F., and Parfitt, G.D. (1987). Interfacial phenomena in apolar media. Surfactant Science Series *21*, 140-160.

Eldridge, J.H., Meulbroek, J.A., Staas, J.K., Tice, T.R., and Gilley, R.M. (1989). Vaccine-containing biodegradable microspheres specifically enter the gut-associated lymphoid tissue following oral administration and induce a disseminated mucosal immune response. Adv Exp Med Biol *251*, 191-202.

Esnouf, R., Ren, J., Ross, C., Jones, Y., Stammers, D., and Stuart, D. (1995). Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors. Nat Struct Biol *2*, 303-308.

Esnouf, R.M., Ren, J., Garman, E.F., Somers, D.O., Ross, C.K., Jones, E.Y., Stammers, D.K., and Stuart, D.I. (1998). Continuous and discontinuous changes in the unit cell of HIV-1 reverse transcriptase crystals on dehydration. Acta Cryst *54*, 938-953.

Felts, A.K., Harano, Y., Gallicchio, E., and Levy, R.M. (2004). Free energy surfaces of beta-hairpin and alpha-helical peptides generated by replica exchange molecular dynamics with the AGBNP implicit solvent model. Proteins *56*, 310-321.

Feng, B.Y., Shelat, A., Doman, T.N., Guy, R.K., and Shoichet, B.K. (2005). High-throughput assays for promiscuous inhibitors. Nature chemical biology *1*, 146-148.

Feng, B.Y., and Shoichet, B.K. (2006). A detergent-based assay for the detection of promiscuous inhibitors. Nature protocols *1*, 550-553.

Forssell, H. (1988). Gastric mucosal defence mechanisms: a brief review. Scand J Gastroenterol Suppl *155*, 23-28.

Fraaije, J.G.E.M., van Vlimmeren, B.A.C., Maurits, N.M., and Postma, M. (1997). The dynamic mean-field density function method and its application to the mesoscopic dynamics of quenched block copolymer melts. J Chem Phys *106*, 4260-4269.

Frenkel, Y.V., Clark, A.D., Jr., Das, K., Wang, Y.H., Lewi, P.J., Janssen, P.A., and Arnold, E. (2005). Concentration and pH dependent aggregation of hydrophobic drug molecules and relevance to oral bioavailability. J Med Chem *48*, 1974-1983.

Frenkel, Y.V., Das, K., Emge, T., Gallicchio, E., Levy, R.J., and Arnold, E. (in preparation). Small molecule crystal structures of TMC278 reveal the role of flexibility in free and bound states. .

Frenkel, Y.V., Gallicchio, E., Das, K., Levy, R.J., and Arnold, E. (submitted). Molecular dynamics study of non-nucleoside HIV-1 RT inhibitor TMC278/rilpivirine aggregates: correlation between amphiphilic properties of the drug and oral bioavailability. Nat Biotech.

Fung, E.N., Chu, I., Li, C., Liu, T., Soares, A., Morrison, R., and Nomeir, A.A. (2003). Higher-throughput screening for Caco-2 permeability utilizing a multiple sprayer liquid chromatography/tandem mass spectrometry system. Rapid Commun Mass Spectrom *17*, 2147-2152.

Gallicchio, E., and Levy, R.M. (2004). AGBNP: an analytic implicit solvent model suitable for molecular dynamics simulations and high-resolution modeling. Journal of computational chemistry *25*, 479-499.

Gallicchio, E., Levy, R.M., and Parashar, M. (2008). Asynchronous replica exchange for molecular simulations. Journal of computational chemistry *29*, 788-794.

Garcia, A.E. (1992). Large-amplitude nonlinear motions in proteins. Physical review letters *68*, 2696-2699.

Grant, B.J., Rodrigues, A.P., ElSawy, K.M., McCammon, J.A., and Caves, L.S. (2006). Bio3d: an R package for the comparative analysis of protein structures. Bioinformatics (Oxford, England) *22*, 2695-2696.

Gruzdev, B., Rakhmanova, A., Doubovskaya, E., Yakovlev, A., Peeters, M., Rinehart, A., de Dier, K., Baede-Van Dijk, P., Parys, W., and van 't Klooster, G. (2003). A randomized, double-blind, placebo-controlled trial of TMC125 as 7-day monotherapy in antiretroviral naive, HIV-1 infected subjects. AIDS *17*, 2487-2494.

Haldar, J., Aswal, V.K., Goyal, P.S., and Bhattacharya, S. (2005). Unusual micellar properties of multiheaded cationic surfactants in the presence of strong charge neutralizing salts. J Colloid Interface Sci *282*, 156-161.

Hammer, S.M., Eron, J.J., Jr., Reiss, P., Schooley, R.T., Thompson, M.A., Walmsley, S., Cahn, P., Fischl, M.A., Gatell, J.M., Hirsch, M.S., *et al.* (2008). Antiretroviral treatment of adult HIV infection: 2008 recommendations of the Int AIDS USA panel. Jamaica *300*, 555-570.

Hill, A., Grossman, H., Cohen, C., Nadler, J., and Peeters, M. (2007). Modelling-based prediction of clinical benefits from etravirine in the TMC125-C223 trial. HIV clinical trials *8*, 68-76.

Hodges, G.M., Carr, E.A., Hazzard, R.A., and Carr, K.E. (1995). Uptake and translocation of microparticles in small intestine. Morphology and quantification of particle distribution. Dig Dis Sci *40*, 967-975.

Hornak, V., and Simmerling, C. (2007). Targeting structural flexibility in HIV-1 protease inhibitor binding. Drug Discov Today *12*, 132-138.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J Mol Graph *14*, 33-38.

Iavicoli, P., Simon-Sorbed, M., and Amabilion, D.B. (2009). Surface aggregate morphology of chiral porphyrins as a function of constitution and amphiphilic nature. New J Chem *33*, 358-365.

Ikomi, F., Hanna, G.K., and Schmid-Schonbein, G.W. (1999). Size- and surface-dependent uptake of colloid particles into the lymphatic system. Lymphology *32*, 90-102.

Jani, P., Halbert, G.W., Langridge, J., and Florence, A.T. (1989). The uptake and translocation of latex nanospheres and microspheres after oral administration to rats. J Pharm Pharmacol *41*, 809-812.

Janssen, P.A.J., Lewi, P., Arnold, E., Daeyaert, F., de Jonge, M., Heeres, J., Koymans, L., Vinkers, H.M., Guillemont, J., Pasquier, E., *et al.* (2005). In search of a novel anti-HIV drug: Multidisciplinary coordination in the discovery of 4-[[4-[[4-[(1E)-2-cyanoethenyl]-2,6-dimethylphenyl]amino]-2-pyrimidinyl]amino]benzonitrile (R278474, Rilpivirine). J Med Chem *in press*.

JMP. Statistical Discovery Software (Version 3.1, SAS Institute Inc., SAS Campus Drive, Cary NC, 1995).

Jorgensen, W.L., Maxwell, D.S., and Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energentics and properties of organic liquids. J Am Chem Soc *118*, 11225-11236.

Junginger, H.E., Hoogstraate, J.A., and Verhoef, J.C. (1999). Recent advances in buccal drug delivery and absorption--in vitro and in vivo studies. J Control Release *62*, 149-159.

Kaminski, G.A., Stern, H.A., Berne, B.J., and Friesner, R.A. (2004). Development of an accurate and robust polarizable molecular mechanics force field from ab initio quantum chemistry. J Phys Chem A *108*, 621-627.

Kaminski, G.A., Stern, H.A., Berne, B.J., Friesner, R.A., Cao, Y.X., Murphy, R.B., Zhou, R., and Halgren, T.A. (2002). Development of a polarizable force field for proteins via ab initio quantum chemistry: first generation model and gas phase tests. J Comp Chem *23*, 1515-1531.

Kipp, J.E. (2004). The role of solid nanoparticle technology in the parenteral delivery of poorly water-soluble drugs. Int J Pharm *284*, 109-122.

Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A., and Steitz, T.A. (1992). Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science *256*, 1783-1790.

Kostewicz, E.S., Brauns, U., Becker, R., and Dressman, J.B. (2002). Forecasting the oral absorption behavior of poorly soluble weak bases using solubility and dissolution studies in biorelevant media. Pharm Res *19*, 345-349.

Kostewicz, E.S., Wunderlich, M., Brauns, U., Becker, R., Bock, T., and Dressman, J.B. (2004). Predicting the precipitation of poorly soluble weak bases upon entry in the small intestine. J Pharm Pharmacol *56*, 43-51.

Kraehenbuhl, J.P., and Neutra, M.R. (2000). Epithelial M cells: differentiation and function. Annu Rev Cell Dev Biol *16*, 301-332.

Lapatto, R., Blundell, T., Hemmings, A., Overington, J., Wilderspin, A., Wood, S., Merson, J.R., Whittle, P.J., Danley, D.E., Geoghegan, K.F., *et al.* (1989a). X-ray analysis of HIV-1 proteinase at 2.7 A resolution confirms structural homology among retroviral enzymes. Nature *342*, 299-302.

Lapatto, R., Hietamaki, A., and Raisanen, J. (1989b). Quantitative trace element analysis of human nails with external beam PIXE. Biological trace element research *19*, 161-170.

Larger, P., Altamura, M., Catalioto, R.M., Giuliani, S., Maggi, C.A., Valenti, C., and Triolo, A. (2002). Simultaneous LC-MS/MS determination of reference pharmaceuticals as a method for the characterization of the Caco-2 cell monolayer absorption properties. Anal Chem *74*, 5273-5281.

Levin, J. (2005). Substantial improvement of oral bioavailability of TMC125 using new tablet formulations in healthy volunteers. Paper presented at: International AIDS Conference (IAS) (Rio de Jeneiro, Brazil).

Levin, J. (2008). Efficacy and safety of TMC278 in treatment-naïve, HIV-infected patients: Week 96 data from TMC278-C204. Paper presented at: 9th International Congress on Drug Therapy in HIV Infection (Glasgow).

Lewi, P., Arnold, E., Andries, K., Bohets, H., Borghys, H., Clark, A., Daeyaert, F., Das, K., de Bethune, M.P., de Jonge, M., *et al.* (2004). Correlations between factors determining the pharmacokinetics and antiviral activity of HIV-1 non-nucleoside reverse transcriptase inhibitors of the diaryltriazine and diarylpyrimidine classes of compounds. Drug R D *5*, 245-257.

Lewi, P.J., de Jonge, M., Daeyaert, F., Koymans, L., Vinkers, M., Heeres, J., Janssen, P.A., Arnold, E., Das, K., Clark, A.D., Jr., *et al.* (2003). On the detection of multiple-binding modes of ligands to proteins, from biological, structural, and modeling data. J Comput Aided Mol Des *17*, 129-134.

Li, A.P. (2001). Screening for human ADME/Tox drug properties in drug discovery. Drug Discov Today *6*, 357-366.

Li, A.P. (2003). Advancing technologies for accelerated drug development. Drug Discov Today *8*, 200-202.

Liu, S., Abbondanzieri, E.A., Rausch, J.W., Le Grice, S.F., and Zhuang, X. (2008). Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. Science *322*, 1092-1097.

Ludovici, D.W., Kavash, R.W., Kukla, M.J., Ho, C.Y., Ye, H., De Corte, B.L., Andries, K., de Bethune, M.P., Azijn, H., Pauwels, R., *et al.* (2001a). Evolution of anti-HIV drug candidates. Part 2: Diaryltriazine (DATA) analogues. Bioorg Med Chem Lett *11*, 2229-2234.

Ludovici, D.W., Kukla, M.J., Grous, P.G., Krishnan, S., Andries, K., de Bethune, M.P., Azijn, H., Pauwels, R., De Clercq, E., Arnold, E., *et al.* (2001b). Evolution of anti-HIV drug candidates. Part 1: From alpha-anilinophenylacetamide (alpha-APA) to imidoyl thiourea (ITU). Bioorg Med Chem Lett *11*, 2225-2228.

Luque, I., Todd, M.J., Gomez, J., Semo, N., and Freire, E. (1998). Molecular basis of resistance to HIV-1 protease inhibition: a plausible hypothesis. Biochemistry *37*, 5791-5797.

Mansky, L.M., and Temin, H.M. (1995). Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. Journal of virology *69*, 5087-5094.

Maple, J.R., Cao, Y.X., Damm, W., Halgren, T.A., Kaminski, G.A., Zhang, L.Y., and Friesner, R.A. (2005). A polarizable force field and coninuum solvations methodology for modeling of protein-ligand interactions. J Chem Theory and Comput *1*, 694-715.

Martin, P., Vickrey, J.F., Proteasa, G., Jimenez, Y.L., Wawrzak, Z., Winters, M.A., Merigan, T.C., and Kovari, L.C. (2005). "Wide-open" 1.3 A structure of a multidrug-resistant HIV-1 protease as a drug target. Structure *13*, 1887-1895.

Martinez, M.N., and Amidon, G.L. (2002). A mechanistic approach to understanding the factors affecting drug absorption: a review of fundamentals. J Clin Pharmacol *42*, 620-643.

McGovern, S.L., Caselli, E., Grigorieff, N., and Shoichet, B.K. (2002). A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. J Med Chem *45*, 1712-1722.

McGovern, S.L., Helfand, B.T., Feng, B., and Shoichet, B.K. (2003). A specific mechanism of nonspecific inhibition. J Med Chem *46*, 4265-4272.

Meyer, E.A., Castellano, R.K., and Diederich, F. (2003). Interactions with aromatic rings in chemical and biological recognition. Angrew Chem Int Ed *42*, 1210-1250.

Narang, A.S., Delmarre, D., and Gao, D. (2007). Stable drug encapsulation in micelles and microemulsions. Int J Pharm *345*, 9-25.

Navia, M.A., Fitzgerald, P.M., McKeever, B.M., Leu, C.T., Heimbach, J.C., Herber, W.K., Sigal, I.S., Darke, P.L., and Springer, J.P. (1989). Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. Nature *337*, 615-620.

Neuhoff, S., Ungell, A.L., Zamora, I., and Artursson, P. (2003). pH-dependent bidirectional transport of weakly basic drugs across Caco-2 monolayers: implications for drug-drug interactions. Pharm Res *20*, 1141-1148.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281-296.

Pacheco, L.F., and Carmona-Ribeiro, A.M. (2003). Effects of synthetic lipids on solubilization and colloid stability of hydrophobic drugs. J Colloid Interface Sci *258*, 146-154.

Pauwels, R., Balzarini, J., Baba, M., Snoeck, R., Schols, D., Herdewijn, P., Desmyter, J., and De Clercq, E. (1988). Rapid and automated tetrazolium-based colorimetric assay for the detection of anti-HIV compounds. J Virol Methods *20*, 309-321.

Perryman, A.L., Lin, J.H., and Andrew McCammon, J. (2006). Optimization and computational evaluation of a series of potential active site inhibitors of the V82F/I84V drug-resistant mutant of HIV-1 protease: an application of the relaxed complex method of structure-based drug design. Chemical biology & drug design *67*, 336-345.

Pidgeon, C., Ong, S., Liu, H., Qiu, X., Pidgeon, M., Dantzig, A.H., Munroe, J., Hornback, W.J., Kasher, J.S., Glunz, L., *et al.* (1995). IAM chromatography: an in vitro screen for predicting drug membrane permeability. J Med Chem *38*, 590-594.

Polak, M., Seley, K.L., and Plavec, J. (2004). Conformational properties of shape modified nucleosides--fleximers. J Am Chem Soc *126*, 8159-8166.

Porter, C.J., Pouton, C.W., Cuine, J.F., and Charman, W.N. (2008). Enhancing intestinal drug solubilisation using lipid-based delivery systems. Adv Drug Deliv Rev *60*, 673-691.

Protein Solutions, I. (1994 - 2000). DynaPro-MS800. In Dynamic light scattering Instruments.

Ramasubbu, N., Parthasarathy, R., and Murray-Rust, P. (1986). Angular preferences of intermolecular forces around halogen centers: preferred directions of approach of electrophiles and nucleophiles around carbon-halogen bond. JACS *108*, 4308-4314.

Regev, O., and Zana, R. (1999). Aggregation Behavior of Tyloxapol, a Nonionic Surfactant Oligomer, in Aqueous Solution. J Colloid Interface Sci *210*, 8-17.

Ripley, B.D. (2001). The R project in statistical computing. MSOR Connections *1*, 23-25.

Rosen, M.J. (1989). Surfactants and interfacial phenomena. Second Edition.

Rosenberg, S.H., Spina, K.P., Condon, S.L., Polakowski, J., Yao, Z., Kovar, P., Stein, H.H., Cohen, J., Barlow, J.L., Klinghofer, V., *et al.* (1993a). Studies directed toward the design of orally active renin inhibitors. 2. Development of the efficacious, bioavailable renin inhibitor (2S)-2-benzyl-3- [[(1-methylpiperazin-4-yl)sulfonyl]propionyl]-3-thiazol-4-yl-L-alanine amide of (2S,3R,4S)-2-amino-1-cyclohexyl-3,4-dihydroxy-6-methylheptane (A-72517). J Med Chem *36*, 460-467.

Rosenberg, S.H., Spina, K.P., Woods, K.W., Polakowski, J., Martin, D.L., Yao, Z., Stein, H.H., Cohen, J., Barlow, J.L., Egan, D.A., *et al.* (1993b). Studies directed toward the design of orally active renin inhibitors. 1. Some factors influencing the absorption of small peptides. J Med Chem *36*, 449-459.

Ryjkina, E., Kuhn, H., Rehage, H., Muller, F., and Peggau, J. (2002). Molecular dynamic computer simulations of phase behavior of non-ionic surfactants. Angew Chem Int Ed Engl *41*, 983-986.

Sankatsing, S.U., Weverling, G.J., Peeters, M., van't Klooster, G., Gruzdev, B., Rakhmanova, A., Danner, S.A., Jurriaans, S., Prins, J.M., and Lange, J.M. (2003). TMC125 exerts similar initial antiviral potency as a five-drug, triple class antiretroviral regimen. AIDS *17*, 2623-2627.

Schlossman, M.L., and Tikhonov, A.M. (2008). Molecular ordering and phase behavior of surfactants at water-oil interfaces as probed by X-ray surface scattering. Annu Rev Phys Chem *59*, 153-177.

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

Seley, K.L., Quirk, S., Salim, S., Zhang, L., and Hagos, A. (2003). Unexpected inhibition of S-adenosyl-L-homocysteine hydrolase by a guanosine nucleoside. Bioorg Med Chem Lett *13*, 1985-1988.

Shoichet, B.K. (2004). Virtual screening of chemical libraries. Nature *432*, 862-865.

Shoichet, B.K. (2006). Screening in a spirit haunted world. Drug Discov Today *11*, 607-615.

Sridharan, S., Nicholls, A., and Honig, B. (1992). A new vertex algorithm to calculate solvent accessible surface areas. Biohys *61*, A174.

Stegemann, S., Leveiller, F., Franchi, D., de Jong, H., and Linden, H. (2007). When poor solubility becomes an issue: from early stage to proof of concept. Eur J Pharm Sci *31*, 249-261.

Stern, H.A., Kaminski, G.A., Banks, J.L., Zhou, R., Berne, B.J., and Friesner, R.A. (1999). Fluctuating charge, polarizable dipole, and combined models: parameterization from ab Initio quantum chemistry. The journal of physical chemistry *103*, 4730-4737.

Stern, H.A., Rittner, F., Berne, B.J., and Friesner, R.A. (2001). Combined fluctuating charge and polarizable dipole models: Application to a five-site water potential function. J Chem Phys *115*, 2237-2251.

SuGI tracta, Y., and Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett *314*, 141-151.

Tsujii, K. (1997). Surface activity: principles, phenomena, and applications.

UNAIDS/WHO (2008). 2008 Report on the Global AIDS Epidemic Update. http://wwwunaidsorg/.

Ungell, A.L., Nylander, S., Bergstrand, S., Sjoberg, A., and Lennernas, H. (1998). Membrane transport of drugs in different regions of the intestinal tract of the rat. J Pharm Sci *87*, 360-366.

Vasconcelos, T., Sarmento, B., and Costa, P. (2007). Solid dispersions as strategy to improve oral bioavailability of poor water soluble drugs. Drug Discov Today *12*, 1068-1075.

Verman, I.M., Temple, G.F., Fan, H., and Baltimore, D. (1974). Synthesis by reverse transcriptase of DNA complementary to globin messenger RNA. Basic life sciences *3*, 355-372.

Waterman, K.C., and Sutton, S.C. (2003). A computational model for particle size influence on drug absorption during controlled-release colonic delivery. J Control Release *86*, 293-304.

Weiss, R.A. (1993). How does HIV cause AIDS? Science *260*, 1273-1279.

White, R.E. (2000). High-throughput screening in drug metabolism and pharmacokinetic support of drug discovery. Annu Rev Pharmacol Toxicol *40*, 133-157.

Xia, Q., Radzio, J., Anderson, K.S., and Sluis-Cremer, N. (2007). Probing nonnucleoside inhibitor-induced active-site distortion in HIV-1 reverse transcriptase by transient kinetic analyses. Protein Sci *16*, 1728-1737.

Yu, L.X., Lipka, E., Crison, J.R., and Amidon, G.L. (1996). Transport approaches to the biopharmaceutical design of oral drug delivery systems: prediction of intestinal absorption. Adv Drug Deliv Rev *19*, 359-376.

Curriculum Vitae

Yulia Frenkel

Education:

1997 – 2001 Cook College, Rutgers University, BA in Chemistry
2002 – 2009 Rutgers University Graduate School, PhD in Biochemistry

Honors and Awards:

1997 – 2001	Golden Key National Honor Society
1997 – 2001	Dean's list of outstanding students
2003 – 2004	NIH Molecular Biophysics Training Grant Award
2005	Dr. Stanley and Francine Mandeles Award
2005	NIH Workforce Training Grant Award
2006	NIH Workforce Training Grant Award
2007 – 2008	NIH Workforce Training Grant Award

Publications:

T. J. Emge, A. Agrawal, J. P. Dalessio, G. Dukovic, J. A. Inghrim, K. Janjua, M. Macaluso, L. L. Robertson, T. J. Stiglic, Y. Volovik and M. M. Georgiadis. **2000**. Alaninyltryptophan hydrate, glycyltryptophan dihydrate and tryptophylglycine hydrate. *Acta Cryst*. **C56**:e469-e471.

Oren, D.A., Y. Li, Y. Volovik, T.S. Morris, C. Dharia, K. Das, O. Galperina, R. Gentz, and E. Arnold. **2002**. Structural basis of BLyS receptor recognition. *Nat. Struct. Biol*. **9**:288-292.

Lewi, P.J., E. Arnold, K. Andries, H. Bohets, H. Brghys, A. D.Clark, Jr., F. Daeyaert, K. Das, M.-P. de Béthune, M. de Jonge, J. Heeres, L. Koymans, J. Leempoels, J. Peeters, P. Timmerman, W. Van Den Broeck, F. Vanhoutte, G. van't Klooster, M. Vinkers, Y. Volovik, and P.A. J. Janssen. **2004**. Correlations between factors determing the pharmacokinetics and antiviral activity of HIV-1 non-nucleoside reverse transcriptase inhibitors (NNRTIs) of the diaryitriazine (DATA) and diarylpyrimidine (DAPY) classes of compounds. *Drugs Res. Devel*. **5**:245-257.

Janssen, P.A., P.J. Lewi, E. Arnold, F. Daeyaert, M. de Jonge, J. Heeres, L. Koymans, M. Vinkers, J. Guillemont, E. Pasquier, M. Kukla, D. Ludovici, K. Andries, M.P. de Bethune, R. Pauwels, K. Das, A.D. Clark, Jr., Y.V. Frenkel, S.H. Hughes, B. Medaer, F. De Knaep, H. Bohets, F. De Clerck, A. Lampo, P. Williams, andP. Stoffels. **2005**. In search of a novel anti-HIV drug: multidisciplinary coordination in the discovery of 4-[[4-[[4-[(1E)-2-cyanoethenyl]-2,6-dimethylphenyl]amino]-2- pyrimidinyl]amino]benzonitrile (R278474, rilpivirine). *J Med Chem* 48(**6**):1901-1909.

Frenkel, Y.V., A.D. Clark, Jr., K. Das, Y.H. Wang, P.J. Lewi, P.A. Janssen, and Arnold, E. **2005**. Concentration and pH dependent aggregation of hydrophobic drug molecules and relevance to oral bioavailability. *J Med Chem* 48(**6**):1974-1983.

Das, K., J.D. Bauman, A.D. Clark, Jr, Y.V. Frenkel., P.J. Lewi, A.J. Shatkin, S.H. Hughes, E. Arnold. **2008**. E. High-resolution structures of HIV-1 reverse transcriptase/TMC278 complexes: Strategic flexibility explains potency against resistance mutations. *PNAS*. 105(**5**):1466-71.

Frenkel, Y.V., K. Das, T. Emge, E. Gallicchio, R.J. Levy, and E. Arnold. (in preparation). Small molecule crystal structures of TMC278 reveal the role of flexibility in free and bound states. .

Frenkel, Y.V., E. Gallicchio, K. Das, R.J. Levy, and E. Arnold. (submitted). Molecular dynamics study of non-nucleoside HIV-1 RT inhibitor TMC278/rilpivirine aggregates: correlation between amphiphilic properties of the drug and oral bioavailability. J Med Chem.