

Report on the PhD thesis  
“Interactions of aminoglycoside antibiotics with  
ribosomal RNA and riboswitches”  
by Ms. Marta Kulik

Antibiotic resistance in bacteria is a serious problem that poses a very real threat to human health. Bacterial infections, which used to take numerous lives every year, are now coming back again. After more than six decades of having a successful defense against these infections in the form of a number of antibiotic drugs, we are now in a situation where several bacterial strains have developed resistance to multiple antibiotics, and at the same time the pharmaceutical industry world-wide is not really coming up with replacement therapeutics.

The subject of this thesis – aminoglycoside antibiotics – is therefore highly relevant. Ribosomes, essential for cellular life, are the target of approximately half of the antibiotics currently in clinical use, and the aminoglycosides are one important class of antibiotics, specifically targeting bacterial ribosomes. One strategy, which can be used to develop new antibiotic drugs that are effective against the current resistance mechanisms in pathogenic bacteria, is to build on existing drugs and re-design them so that they retain their antibiotic activity against the resistant strains. This requires a thorough understanding both of how the drugs work in the first place, and also of how some bacteria manage to avoid being killed.

Ms. Kulik addresses these issues in studies of three different RNAs, which all are targets for aminoglycosides: The ribosomal A-site RNA (where the first recognition of the incoming amino-acyl tRNA is made), ribosomal Helix 69 (at the interface between the small and large ribosomal subunits), and a synthetic construct, the N1 riboswitch. The two first are natural antibiotic targets, and the riboswitch is a small RNA that shares some binding site features with the natural targets, which makes it an easy to use model system, even though it is more related to gene regulation than to antibiotic activity per se. Peptide Nucleic Acid (PNA), a non-aminoglycoside, is also used with the Helix 69 target. Both computational (energy minimization and analysis, molecular dynamics simulation) and experimental (spectroscopy, calorimetry) methods are used. This broad combination of techniques is a very strong aspect of the thesis.

Overall the thesis is well-written and clearly presents the biological/biochemical systems, the methodological approaches, and the results. One of the main achievements is the demonstration that a computationally expensive, more refined,

method of treating charge distributions (aspherical distributions) does not provide any significant improvement over simpler and less expensive point charge representations in these highly charged biological macromolecules. This is important to know since computational power is still a limiting factor in many of these studies, and the community therefore always strives to use efficient, yet **reliable**, approximations. A new antibiotic compound was proposed, based on the interaction analysis of a set of aminoglycosides with the RNA target; this compound should be testable, and so is a bona fide prediction – always a very good outcome of a computational study! Guided by interaction involving stably bound water molecules at the drug-RNA interface (a useful, but in itself not new strategy), the author also designed a new, more rigid, variant of aminoglycosides. For Helix 69 a nucleic analog (PNA) that should have inhibitory effect was designed. Finally, the activation of the N1 riboswitch is suggested to be related to its flexibility, which is differentially affected by active and inactive ligands.

All these drugs bind non-covalently to their targets, and the key thermodynamic quantity is the **free energy of binding ( $\Delta G$ )**. The computational analyses in the thesis are mostly concerned with either structural aspects (*e.g.* number of hydrogen bonds) or (electrostatic) interaction energies – which are not the same thing as  $\Delta G$ . Yet, there are well-developed methods for computing relative  $\Delta G$ -values (the difference in binding free energy between two compounds or two targets) from simulations. It would have been of interest to see this type of free energy perturbation analysis in some of the systems, or at least to have a discussion of pros and cons of the free energy perturbation method.

The limitations of the methods used are not discussed in detail, and likewise estimates of convergence and errors are not presented in most cases. Is the length (100 ns) of the molecular dynamics simulations sufficient to obtain stable and converged results? This could easily be checked by running several independent 100 ns simulations of the same system.

Below I give some specific comments for each chapter in the thesis.

### **Chapter 1.**

Why is the problem of antibiotic resistance especially serious with Gram negative bacteria?

Is there really a need for > 50% G in the homopurine strand of the duplex part in a triplex (page 12)? In DNA this is not necessary.

Stacking interactions do exist even between rings that are not charged (the nucleobases are not charged).

## Chapter 2.

The precision in the structures of these molecules is limited (resolutions of 2.5Å or worse), and they are not rigid, so the assumption that electrostatic interaction energies can be improved by using an aspherical charge distribution for the atoms seems to be a bit shaky. The issue is further complicated by the energy minimization, which uses a different charge model, and which in itself may not give the optimal representation of the real system which is flexible and may visit several local conformations; a charged side chain may easily swing several Å from its position in the X-ray or NMR structure, and this changes its electrostatic interactions (especially *in vacuo*) by tens of kcal/mol. The very large values of the electrostatic interaction energies make it difficult to trust their predictive value – would they really all be scaled by the same factor if solvent screening was taken into account? There are also interactions within the RNA that have not been accounted for.

In equations 2.6 and 2.7, why is the sum from 1 to  $N/2$ ? Does this mean that  $N$  has to be even?

There is a typographical error in equation 2.13, there should not be a squaring of the dihedral term.

Please explain the columns and rows of the matrix **D** in equation 2.18.

Why not simply use the microcanonical ensemble (page 32)?

Why is the method with fixed partial charges “the only methods that allows us to calculate the forces action on the nuclei” (page 37). How about polarizable models, or lattice sum methods like Particle Mesh Ewald summation? What is in general the computational difficulty with electrostatic interactions?

Is the absorbance  $A$  the amount of absorbed light (page 39)? And is the expression given really the Lambert-Beer law?

## Chapter 3.

Why were the 5'-phosphates removed? This interferes with the estimate of interaction with G1403 on page 50.

The AMBER force field contains parameters, including partial charges, for the RNA residues, so the AM1-BCC procedure mentioned on page 46 presumably relates to the ligands. How were the rest of the parameters for the ligands obtained?

What is the “conversion factor” on page 47? Is this something particular for AMBER?

In the legend of figure 3.2 it should be vdW **energies** (not forces).

If charges dominate the interactions it is unclear why mutations of non-charged bases should have strong effects (Table 3.5).

Why should a strongly bound water molecule not be used to guide the introduction of a modification to the ligand (page 71)?

Biomolecular crystals usually are quite wet (50% or more water), so the crystal structures may not be solidly packed states (page 75).

#### Chapter 4.

Is the PNA binding to Helix 69 an example of antisense binding? This is not a nucleic acid that is being read in a transcription or translation process (where “sense” makes sense). And are the binding modes of neomycin and PNA related?

Why does mismatched or random PNA reduce  $T_m$  of the hairpin?

In figure 4.6, how does the excess PNA influence the CD spectrum?

The IC50 values, free energies and MICs (pages 96-97) are a bit confusing. With neomycin and PNA the IC50s differ, but free energies are similar, and with tetracycline and PNA the IC50s are similar but the MICs differ?

How does PNA adhesion to the RNA-PNA duplex occur?

What does “one-fold lower (*i.e.* “1 times lower”) mean on page 96?

#### Chapter 5.

There is significantly more simulation time in the REMD (32x 100 ns) simulation compared to the standard MD simulation (1 x 100 ns). How can one know that the REMD is more efficient? What would happen if you were to run 32 x 100 ns (or 8 x 400 ns) standard MD?

Force field validation is always a problem, generally due to lack of useful experimental data, and you have a handful of NMR-derived interproton distances. How much do you trust the force field here?

The standard deviations in figures 5.7 and 5.8 could be replaced by the standard error of the mean, which is more informative regarding the precision of the average value.

The number of hydrogen bonds between ligand and RNA is surprisingly low, less than 50% (page 111) – comment?

How does the motion of A17 act as a signal (page 113)?

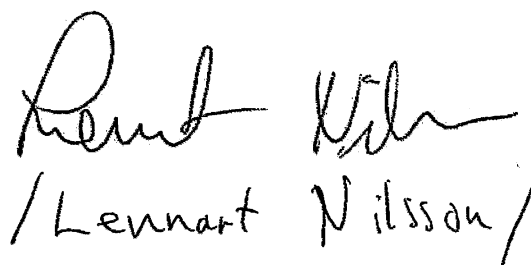
How can ring IV affect the motions of ring I (figure 5.10)? Are the principal components just for the motions of ring I alone?

Can NMR be performed on a frozen solution?

Is it the RNA shape or its sequence that matters? The message is different regarding this on pages 117 (shape) and 118 (sequence).

**Final evaluation.**

This thesis deals with important biological/biochemistry problems using a wide set of appropriate methods, and the results of these investigations are clearly presented both in the thesis and in several publications in international journals, as well as at a number of conferences. I conclude that the thesis clearly meets the criteria for the PhD degree.

  
/Lennart Nilsson/

# Raport nt. pracy doktorskiej

## “Interactions of aminoglycoside antibiotics with ribosomal RNA and riboswitches”

### Pani Marty Kulik

Oporność na antybiotyki u bakterii stanowi poważny problem, będący realnym zagrożeniem dla zdrowia ludzkiego. Infekcje bakteryjne, które każdego roku pozbawiały życia setki istnień, znowu powracają. Po więcej niż sześciu dekadach udanej obrony przeciwko tym infekcjom za pomocą antybiotyków, jesteśmy w sytuacji, w której niektóre szczepy bakterii rozwinęły oporność wobec wielu antybiotyków i jednocześnie światowy przemysł farmaceutyczny nie dostarcza zastępczych leków.

Tematyka tej pracy – antybiotyki aminoglikozydowe – ma zatem szczególne znaczenie. Rybosomy, niezbędne dla życia komórkowego, są celem blisko połowy antybiotyków obecnie stosowanych w praktyce klinicznej a aminoglikozydy są jedną z ważnych klas antybiotyków, specyficznie celujących w rybosomy bakteryjne. Jedną ze strategii, która może zostać wykorzystana do rozwijania nowych antybiotyków skutecznych w walce przeciwko obecnym mechanizmom oporności u bakterii patogennych, polega na użyciu istniejących leków i ich przeprojektowaniu tak, by zachowały swoją aktywność antybakteryjną przeciwko opornym szczepom. Wymaga to głębokiego zrozumienia zarówno tego, jak działają leki, jak i tego, jak niektóre bakterie unikają śmierci.

Pani Kulik odnosi się do tych problemów w badaniach trzech różnych RNA, które są celami dla aminoglikozydów: rybosomalnego miejsca A (gdzie następuje pierwsze rozpoznanie przychodzących amino-acylowanych tRNA), rybosomalnej Helisy 69 (na styku pomiędzy małą a dużą podjednostką rybosomu), oraz syntetycznym konstruktym, ryboprzełącznikiem N1. Dwa pierwsze RNA są naturalnymi celami antybiotyków, natomiast ryboprzełącznik jest małym RNA, które posiada wspólne cechy miejsc wiązania z celami naturalnymi, co powoduje, że jest on łatwym do wykorzystania systemem modelowym, choć jednocześnie bardziej powiązany z regulacją genetyczną niż z aktywnością antybakteryjną *per se*. Peptydowy kwas nukleinowy (PNA), niebędący aminoglikozydem, jest także użyty z Helisą 69 jako celem. Używane są zarówno metody obliczeniowe (minimalizacja energii i analiza, symulacje dynamiki molekularnej), jak i eksperymentalne (spektroskopia, kalorymetria). Ta szeroka kombinacja technika stanowi silną stronę rozprawy.

Z ogólnego punktu widzenia, rozprawa jest dobrze napisana i jasno prezentuje biologiczne/biochemiczne układy, podejścia metodologiczne i wyniki. Jednym z głównych osiągnięć jest zademonstrowanie, że obliczeniowo wymagające, bardziej udokładnione metody traktowania rozkładów ładunku,

podejście asferyczne) nie przynoszą znaczącej poprawy wobec prostszych i mniej wymagających reprezentacji ładunków punktowych w tak wysoko naładowanych biologicznych systemach makromolekuł. Jest to ważna informacja, gdyż moc obliczeniowa jest nadal czynnikiem ograniczającym wiele tego typu badań, a w środowisku kładzie się nacisk, aby używać nie tylko wydajnych ale i **wiarygodnych** przybliżeń. Opierając się na analizie oddziaływań szeregu aminoglikozydów z ich celem RNA, zaproponowany został nowy związek antybiotyczny; związek ten powinien być przetestowany, i takie jest przewidywanie *bona fide* – jest to zawsze dobry wynik pracy obliczeniowej! Na podstawie oddziaływań angażujących cząsteczki wody, stabilnie związane na pograniczu leku i RNA (użyteczna, ale nie nowa strategia), autorka zaprojektowała nowy, bardziej sztywny wariant aminoglikozydu. Dla Helisy 69 został zaprojektowany odpowiednik kwasu nukleinowego (PNA), który powinien posiadać właściwości inhibitora. Wreszcie, aktywacja ryboprzełącznika N1 została powiązana z jego elastycznością, na którą różny wpływ mają aktywne i nieaktywne ligandy.

Wszystkie te leki wiążą się niekowalencyjnie do swoich celów, a kluczową wielkością termodynamiczną jest **energia swobodna wiązania ( $\Delta G$ )**. Analiza obliczeniowa w pracy jest głównie związana albo z aspektami strukturalnymi (np. liczbą wiązań wodorowych) albo z (elektrostatyczną) energią oddziaływania – nie jest to to samo, co  $\Delta G$ . Jednakże, istnieją dobrze opracowane metody obliczania względnych wartości  $\Delta G$  (różnicy swobodnych energii wiązania pomiędzy dwoma związkami bądź dwoma celami) z symulacji. Byłoby interesujące zobaczyć w przypadku niektórych systemów analizę wykonaną metodą zaburzeń energii swobodnej, lub przynajmniej pokazać dyskusję wad i zalet tej metody.

Ograniczenia użytych metod nie zostały szczegółowo przedyskutowane, tak jak i oszacowania zbieżności i błędów w większości przypadków nie zostały pokazane. Czy długość (100 ns) symulacji dynamiki molekularnej jest wystarczająca do otrzymania stabilnych i zbieżnych wyników? Może to zostać łatwo sprawdzone poprzez wykonanie kilku niezależnych 100 ns symulacji tego samego systemu.

Poniżej zamieszczam kilka szczegółowych komentarzy do każdego rozdziału pracy.

## Rozdział 1.

Dlaczego problem oporności na antybiotyki jest szczególnie poważny w przypadku bakterii Gram-ujemnych?

Czy rzeczywiście potrzebne jest > 50% G w nici zawierającej puryny, należącej do dupleksu, w tripleksach (strona 12)? W przypadku DNA nie jest to konieczne.

Oddziaływania warstwowe istnieją nawet pomiędzy pierścieniami, które nie są naładowane (zasady w kwasach nukleinowych nie są naładowane).

## Rozdział 2.

Precyzja struktur tych cząsteczek jest ograniczona (rozdzielczość rzędu 2.5 Å lub gorsza), i nie są one sztywne, zatem założenie, że oddziaływania elektrostatyczne mogą być poprawione poprzez użycie asferycznego rozkładu ładunku dla atomów wydaje się być trochę na wyrost. Problem ten jest dodatkowo komplikowany przez zastosowanie minimalizacji energii, która wykorzystuje inny model ładunku, i sama w sobie może nie dawać optymalnej reprezentacji rzeczywistego systemu, który jest elastyczny i może odwiedzać kilka lokalnych konformacji; naładowany łańcuch boczny może łatwo przeskakiwać kilka Å od swojego położenia w strukturze krystalograficznej lub NMR, a zmiany jego oddziaływań elektrostatycznych (szczególnie *in vacuo*) mogą się różnić o dziesiątki kcal/mol. Tak duże wartości elektrostatycznej energii oddziaływania sprawiają, że trudno ufać w ich moc przewidywania – czy rzeczywiście będą się skalowały tak samo po uwzględnieniu ekranowania od rozpuszczalnika?

Także oddziaływania w obrębie RNA nie zostały wzięte pod uwagę.

Dlaczego w równaniach 2.6 i 2.7 występuje suma od 1 do  $N/2$ ? Czy to oznacza, że  $N$  powinno być parzyste?

W równaniu 2.13 jest błąd typograficzny, nie powinno być podnoszenia do kwadratu członu kąta dwuściennego.

Proszę wyjaśnić kolumny i rzędy macierzy **D** w równaniu 2.18. Czemu by nie użyć po prostu zespołu mikrokanonicznego (strona 32)?

Czemu metoda z ustalonymi ładunkami cząstkowymi jest „jedyną metodą, która pozwala na obliczenie sił działających na jądro” (strona 37). Co z modelami uwzględniającymi polaryzację, albo z metodami sum sieciowych, np. Particle Mesh Ewald? Jaka jest w ogólności trudność obliczeniowa, związana z oddziaływaniami elektrostatycznymi?

Czy absorbancja  $A$  jest ilością zaabsorbowanego światła (strona 39)? I czy podane wyrażenie jest rzeczywiście prawem Lamberta-Beera?

## Rozdział 3.

Czemu grupy fosforanowe na końcach 5' zostały usunięte? To zakłóca oszacowanie oddziaływania z G1403 na stronie 50.

Pole siłowe AMBER zawiera parametry, włączając w to ładunki cząstkowe, dla cząsteczek RNA, zatem procedura AM1-BCC, wspomniana na stronie 47, odnosi się zapewne do ligandów. Jak reszta parametrów została wyznaczona dla ligandów?

Co to jest „współczynnik konwersji” na stronie 47? Czy jest to coś specjalnego dla AMBERA?

W legendzie rysunku 3.2 powinny być **energje** vdW (nie siły).



Jeśli ładunek dominuje w oddziaływaniach, nie jest jasne, czemu mutacje nienaładowanych zasad miałyby mieć silny efekt (Tabela 3.5).

Czemu silnie związana cząsteczka wody nie może być użyta do zaproponowania modyfikacji ligandu (strona 71)?

Kryształy biomolekularne są zwykle dosyć mokre (50% lub więcej wody), zatem struktury krystalograficzne nie mogą być silnie upakowane (strona 75).

#### Rozdział 4.

Czy wiązanie PNA do Helisy 69 jest przykładem wiązania antysensownego? Nie jest to kwas nukleinowy odczytywany w transkrypcji w procesie translacji (gdzie „sens” ma sens). Czy sposoby wiązania neomycyny i PNA są ze sobą powiązane?

Czemu niekomplementarna lub losowa sekwencja PNA obniża  $T_m$  spinki?

W jaki sposób nadmiar PNA wpływa na widmo CD na rysunku 4.6?

Wartości IC50, energie swobodne i MIC (strony 96-97) są nieco dezorientujące. W przypadku neomycyny i PNA wartości IC50 się różnią, ale energie swobodne są podobne, a w przypadku tetracykliny i PNA, IC50 są podobne, ale MIC są różne? Jak zachodzi adhezja PNA do kompleksu RNA-PNA?

Co oznacza “o rząd niższy (tzn. “1 raz niższy”) na stronie 96?

#### Rozdział 5.

Czas symulacji jest znacząco dłuższy w przypadku REMD (32x 100 ns) w porównaniu do standardowej symulacji MD (1 x 100 ns). Jak można rozpoznać, że REMD jest bardziej wydajna? Co by się stało, gdyby przeprowadzić 32 x 100 ns (albo 8 x 400 ns) standardowej MD?

Sprawdzenie pola siłowego jest zawsze problemem, generalnie z powodu braku użytecznych danych eksperymentalnych, a tu masz garść odległości między protonami z danych NMR. Jak bardzo tutaj ufasz polu siłowemu?

Odchylenia standardowe na rysunkach 5.7 i 5.8 mogą być zastąpione standardowymi błędami średniej, które wnoszą więcej informacji, dotyczących precyzji wartości średniej.

Liczba wiązań wodorowych pomiędzy ligandami a RNA jest zaskakująco niska, poniżej 50% (strona 111) – komentarz?

W jaki sposób ruch A17 działa jako sygnał (strona 113)?

Jak pierścień IV wpływa na ruch pierścienia I (rysunek 5.10)? Czy główne składowe są podane tylko dla pierścienia I?

Czy NMR może być wykonane dla zamrożonego roztworu?

Czy liczy się kształt RNA czy sekwencja? Podana jest różna informacja, biorąc pod uwagę strony 117 (kształt) i 118 (sekwencja).

### **Ostateczna ocena.**

Ta praca dotyczy ważnych biologicznie/biochemicznie problemów z użyciem szerokiego zestawu odpowiednich metod, a wyniki badań są przedstawione jasno zarówno w pracy, jak i w kilku publikacjach w międzynarodowych czasopismach, oraz na licznych konferencjach. Podsumowując, praca wyraźnie spełnia kryteria jako podstawa do nadania stopnia doktora.