**Crystal Molecular Dynamics**

**by**

**Paweł A. Janowski**

**A dissertation submitted to the**

**Graduate School of New Brunswick**

**Rutgers, the State University of New Jersey**

**in partial fulfillment of the requirements**

**for the degree of**

**Doctor of Philosophy**

**Graduate Program in Chemistry and Chemical Biology**

**and**

**Graduate Program in Computational Biology and Molecular Biophysics**

**Written under the direction of**

**Prof. David A. Case**

**and**

**Prof. Darrin M. York**

**and approved by**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**New Brunswick, New Jersey**

**August, 2015**

# Abstract of the dissertation

**Improved Molecular Dynamics and Macromolecular Crystallography through Simulations of Biomolecular Crystals**

**By Paweł A. Janowski**

**Dissertation Directors**

**David A. Case Ph.D and Darrin M. York Ph.D**

We present a broad effort at the development of crystal simulation methodology and its application to benefit both macromolecular crystallography and molecular dynamics methods. Crystallography is the current method of choice for structural determination of biomolecules, but it is hampered by the inherently time and space averaged nature of the experiment as well as methodological limitations that do not sufficiently account for the heterogeneous and dynamic nature of crystals. Molecular dynamics has proven itself as a method capable of probing the physics and chemistry of biomolecules on an atomic scale, but requires continued development of the underlying force field parameters to more accurately reproduce observables. Our effort has focused on developing the framework for molecular dynamics simulations of biomolecular crystals. We first present our methodology for performing crystal simulations and show how it is applied first to simple peptide crystals and then to increasingly complex biomolecular systems. Next we demonstrate the utility of crystal simulations for validation of molecular dynamics methods through two case studies of the biophysics of enzyme reactions. Finally we demonstrate the improvement to crystallographic methods that can be gained by incorporating molecular dynamics methods. Our work is of great benefit to both the molecular dynamics and macromolecular crystallography communities and proposes specific approaches to integrate the two fields for the benefit of both.

# Acknowledgements

I thank the Lord for the beautiful gift of His Creation, for the flowers, the mountains, the creatures and the fascinating molecular mechanisms of life that testify to His love. Thank you for the gift of reason, free will and desire of the Good, the True and the Beautiful that drives us in all our pursuits. May we one day attain the fullness of Love.

I thank Maria my wife who accompanied me throughout this path in person and in spirit. Thank you for your care, your patience, your inspiration. We have truly walked the past five years together in friendship. May we walk many more!

I thank my parents, Jolanta and Andrew, for their love in raising me to be the person I am. Thank you for teaching me what is important in life and what is not. I am forever indebted in the bond of filial love.

I thank my family and my friends for all your kindess, your love, all the good times spent together that will forever form part of the treasure of my memories. Without those wonderful times shared together, I would have been hard pressed to keep my sanity along the way.

I thank all my colleagues, my lab mates, my teachers, all the wonderful scientists I have had the pleasure to work with. Thank you for showing me how exciting science is!

I thank all of the good people I have met. Each encounter with each one of you enriches both of us. Only in relationship with others does man become truly man. Never stop being who you are. I love you all.

Finally, and in a special way, I thank my advisors, Prof. David Case and Prof. Darrin York. You have challenged and inspired me constantly and at the same time been caring guides along the path. I have learned so much from you, not just about science but about what it means to be good human beings. I will remain forever grateful.

To the loving memory of my Father who would have wanted to be here but is even closer than we can imagine.

Contents

[Abstract of the dissertation ii](#_Toc423532554)

[Acknowledgements iv](#_Toc423532555)

[Contents vi](#_Toc423532556)

[Abbreviations used vii](#_Toc423532557)

[Chapter 1. Introduction 1](#_Toc423532558)

[1.1. Biomolecular Crystallography 2](#_Toc423532559)

[1.2. Molecular Dynamics 2](#_Toc423532560)

[1.3. Goals and overview 2](#_Toc423532561)

[Chapter 2. Developing molecular dynamics of crystals. 3](#_Toc423532562)

[2.1. Fav8 1 3](#_Toc423532563)

[2.2. Fav8 2 3](#_Toc423532564)

[2.3. 4lzt 3](#_Toc423532565)

[2.4. DNA/RNA 3](#_Toc423532566)

[Chapter 3. Applications of molecular dynamics of crystals 3](#_Toc423532567)

[3.1. Hairpin 3](#_Toc423532568)

[3.2. RnaseA 3](#_Toc423532569)

[Chapter 4. Improved crystallographic methods through crystal molecular dynamics 3](#_Toc423532570)

[4.1. AFITT 3](#_Toc423532571)

[4.2. Phenix-Amber 4](#_Toc423532572)

# Abbreviations used

MD – Molecular dynamics;

BX – Biomolecular crystallography;

# Introduction

When, during my initial visit to Rutgers, Prof. David Case first mentioned the idea of improving crystallography through molecular dynamics of crystals, I felt a tinge of excitement. I had studied crystallography for two semesters during my undergraduate coursework at Jagiellonian University in Krakow. Lectures were eloquently delivered by one of the best teachers I’ve ever had, Prof. Krzysztof Lewiński. However, despite all my effort I could not grasp the essence of how a seemingly random pattern of dots on a sheet of paper could be turned into a three dimensional model of a biomolecule. I liked crystallography, but I also respected it and I feared it because I felt like there was something powerfully beautiful and mysterious about it. So when Dr. Case floated this idea of molecular dynamics of crystals I was excited: here was a chance to make up for my previous failing, to finally come to understand crystallography or to die trying. And to do that by using the molecular dynamics that I wanted to focus my Ph.D. studies on… it was the perfect project.

Thus I have happily spent the last five years focused on our effort to simulate biomolecular crystals with molecular dynamics. The original question we asked ourselves was simple: what can we learn from molecular dynamics of crystals? And this was quickly reformulated into the following four overarching questions that form the focus of this work:

1. What is the best way to carry out molecular dynamics of biomolecular crystals?
2. How can we use crystal simulations to improve molecular dynamics methods?
3. How can we use crystal simulations to improve crystallography methods?
4. What can we learn about real crystals from our simulations of crystals?

What follows is a brief introduction to the methods of crystallography and molecular dynamics, with special emphasis on aspects that relate directly to our work. We then discuss the goals and specific aims of this research and present the general organization of the dissertation before moving on to a presentation of the work in subsequent chapters.

## Biomolecular Crystallography

### What is crystallography

Blah blah

### Theory

Blah blah

### Refinement

Blah blah

Crystallography is a biophysical technique used to probe the three-dimensional distribution of atoms in molecules. As the name suggests, requires crystals … Proteins known to crysallize for long time… Crystallography based on diffraction on the regular array… Light … In … Max Laue… diffract on crystal but need wavelength: x-rays. .. Later Braggs the law rules where diffraction spots form… The first protein crystal … In all …. Nobel prizes.

The raw experimental data obtained in a crystallography experiment is a diffraction patter (fig). This is obtained as the x-rays scatter on the electron clouds of the atoms. In some cases con structive interference… This contains essentially two pieces of information: location of spots and relative strength (brightness) of the spots. We now examine each in detail.

The location of spots provides information on the crystal lattice itself (….). Interpreattion based on Bragg’s Law. In 1-d grating you get lines. In 2-d you get dots. In 3-d you get dots located on sphere. Where dots form is determined by where constructive interference (waves on waves). This is the Braggs law. Must arrive so that crests match up. When? When the waves differ by integer. …Reciprocal lattice?

Thus from spots we know the lattice parameters. The other piece of information is relative intensity of the spots. This tells us about the contents of each unit cell, i.e. the structure of the molecules. Figure… Thus if atoms only on the plane, strong spot, if on/off, no spot, and the distribution in between something in between. But this is very complex as the distribution affects all the spots. As it turns out the relationship between the intensity of the spots and the electron density can be formulated via a well know mathematical relation known as the Fourier Transform. For instructive purposes, 1D. The intensity of a spot is … This is the FT…

Now let us examine. What do we need to calculate the the electron density: amplitudes and phases. But what do we have, amplitudes. This is the phase problem. Many techniques have been devised and not go in here. Suffice it to say that a general estimate of phases is sufficient to continue. Majority of structures is based on molecular replacement where an estimate of the phase is obtained using the model of a similar molecule. An estimate model…

Once an estimate model is obtained, the next stage is refinement which is most pertinent here. How can we continue once we have that? FT forward need electron density at every point on a 3D grid. This can be calculated from the distribution of atoms using the …. FT of ED gives us the amplitudes and phases. We don’t know about the phases. But we know the amplitudes. We can compare the two. Usually for this comparison a statistic called the R-factor is used. Smaller the R, the smaller the sum of differences between the observed and calculated amplitudes (Fobs, Fcalc). If not perfect, we can move the atoms around and recalculate the R-factor and keep moving until better.

We could do this by hand and it might work or we might be at it forever. Mathematical schemes to minimize. In the basic approach minimize the sum of square difference between the amplitudes (F(xyz)). Not well defined… parameter to observed ratio Alternatively we can use a maximum likelihood formulation. Advantage of allowing a Bayesian treatment. Minus log and minizimize is same as finding the max. The now becomes. Why prior? Because increases the observables… adds additional constraints, lowers the search space of the optimization algorithm.

Prior… most programs use EH. What is EH… But originally MD… Moved away from… (Maybe sec 3).

In practice more complicated. XYZ refinement not robust enough to find the optimal location- stuck in local minima. Rounds of refinement, manual rebuilding. Second many more parameters. Fluctuations modelled as B-factor (formula)… refine B-factors as isotropic or anisotropic. TLS parameters. Occupancy and alternate conformations. Bulk solvent and anisotropic scaling factor. Macrocycles… in Phenix… In any case one ends up with a 3D model of the locations of the atoms in the crystal.

## ***Molecular Dynamics***

### What is Molcular Dynamics?

Blah blah

### Theory

Blah blah

### Molecular Dynamics of Crystals

Molecular dynamics is... F=ma, U=… etc. Standard force field used in Amber… Each term means…

Additional stuff, such as thermostat, shake, etc.

Standard simulation is done in solvated box, not crystal. Usually insert, remove, equilibrate…

MD has shown itself to be remarkably useful and successful. In the core just a simplified model. So much not modelled, but … examples….

## Goals and overview

Both methods exteremely valuable but also suffer from limiations. These limitations can be overcome at least in part through MD of crystals. Let’s look at each one.

Crystallography: First sources of error and noise. Sometimes so high (low res) that indeterminate. Simulations if reliable could tell us more to help resolve. Second, time and space average… End up with single static view. Myopic because one best rep view of the average. In fact crystals move (dynamic) and heterogenous. Recent efforts by several group aimed at resolving this. Ensemble refinement. Networks stuff from Fraser. Diffuse scattering from … Move and insights about functions. Also insights about crytals. By simulating over multiple copies (space) and ns (time) we can undo the averaging and get time-space resolved glimpse. A more comprehensive view of the crystal. Also, information about crystals: solvent distribution, etc lead to better refinement techniques. Finally, in itself a better set of priors.

MD is good but usually run as solvated box. Good sense but drawback that not comparing against experimental… Validation from crystals. Look directly at experimentals such as structure, fluctuations, electron density and amplitudes… Can check how well it’s doing and modify ff…

Organization: Part II developing methodology for simulating crystals. Apply to larger and conclusions about MD… Part III crystal syms applied to scientific investigation : two cases where used to validate Part III md of crysals applied to improve cyrstallograpjhy methods… Copyrights…

# Developing molecular dynamics of crystals.

## Fav8 1

Blah blah

## Fav8 2

Blah blah

## 4lzt

Blah blah

## DNA/RNA

Blah blah

# Applications of molecular dynamics of crystals

## Hairpin

Blah blah

## RnaseA

Blah blah

# Improved crystallographic methods through crystal molecular dynamics

## AFITT

Blah blah

## Phenix-Amber

Blah blah