# Modeling Metal Protein Complexes from Experimental Extended X-ray Absorption Fine Structure using Evolutionary Algorithms

Collin Price
Brock University
St. Catharines, Ontario
Email: cp06vz@brocku.ca

Sheridan Houghten
Brock University
St. Catharines, Ontario
Email: houghten@brocku.ca

Sergey Vassiliev
Brock University
St. Catharines, Ontario
Email: svassiliev@brocku.ca

Doug Bruce
Brock University
St. Catharines, Ontario
Email: dbruce@brocku.ca

***Abstract*—Experimental *extended x-ray absorption fine structure* (EXAFS) spectra carry information about the chemical structure of metal protein complexes. However, predicting the structure of such complexes from EXAFS spectra is not a simple task. Currently methods such as Monte Carlo Optimization or simulated annealing are used in structure refinement of EXAFS. These methods have proved somewhat successful in structure refinement but have not been successful in finding the global minima. Based on the success of using evolutionary algorithms to overcome local minima issues in other domains, we propose an approach using a genetic algorithm (GA) to better predict the structure of metal protein complexes.**

***Keywords*—Molecular Structure, Evolutionary Algorithms, Representation, Recentering-Restarting, Particle Swarm Optimization, Extended X-ray Absorption Fine Structure,**

## I. Introduction

The goal of this work is to explore different types of evolutionary algorithms in an effort to find the exact molecular structure of a molecule given experimental *extended X-ray absorption fine structure* (EXAFS) spectra. EXAFS experiments are an effective way of determining a molecule's molecular structure. EXAFS experiments are unique in that they can be used on a variety of materials, including gas and liquids.

For the experiments a non-protein chemical compound found within proteins is used. The oxygen-evolving complex (OEC) is required to catalyze the oxidation of water to produce dioxygen. There is a high quality experimental EXAFS spectra of OEC in $S_1$-state that will be used to create a comparative measure against the results. The goal is to create a calculated molecular structure that closely resembles the experimental EXAFS spectra. Finding a good molecular structure for the OEC is important in understanding the catalytic cycle and the development of biomimetics [1].

Currently such methods as computational chemistry [2], quantum mechanics [2] [3], and conjugate gradient optimization [1] are being used to determine the molecular structure of compounds. The problem with these approaches is that they are converging on local minima. A major difficulty with molecular structure refinement is that there are too many degrees of freedom involved. With respect to the OEC there are 79 degrees of freedom, corresponding to the 79 atoms in the OEC. See Figure VI for a breakdown of the different atoms in the OEC. A genetic algorithm (GA) [4] will be used in an attempt to better explore the search space.

## II. Structure Refinement Problem

### A. EXAFS

The following overview is based on information contained in Matthew Newvilles Fundamentals of XAFS (2004) [5]. X-Ray absorption fine structure (XAFS) is a method used to measure the absorption coefficient of a material as a function of energy. X-rays are part of the electromagnetic spectrum with wavelengths ranging from 25Å to 0.25Å. All atoms resonate at a specific wavelength. The x-ray is tuned to have the same wavelength as the target atom. A photon from an x-ray is absorbed by an electron in a tightly bound quantum core level of an atom. Absorption only takes place if the binding energy of the core level is less than the energy of the x-ray photon. At the time of absorption a core electron moves to an empty outer shell and another electron moves in to take its place. Eventually the affected electrons decay to their original state. During this time fluorescence energies are emitted that characterize a specific atom.

The absorption coefficients measured after the initial absorption are referred to as the EXAFS. During the decay of the electrons to their original state, oscillations occur in the measure of the absorption coefficient. The different frequencies found within the oscillations correspond to different near-neighbour coordination shells, which can be described and modeled according to the EXAFS equation. From the oscillations the number of neighbouring atoms, the distances to the neighbouring atoms, and the disorder in the neighbour distances can be determined.

### B. EXAFS Spectra Fitting

EXAFS can be used to identify properties of a molecule, but they do not provide enough detail to determine the atomic structure of a molecule in 3-dimensional space. An EXAFS spectra allows you to identify how far apart atoms are from each other, but does not give enough information to identify their dihedral angles. Fortunately, EXAFS can be used to assist in determining the atomic structure of a molecule. The energy spectra given off by the molecule is unique for its structure, meaning that you can create an atomic structure,

obtain its EXAFS spectra, and compare the results. The hope is that if you create an atomic structure whose spectra closely matches the spectra of an actual model, then there is a high likelihood that the created structure will closely match the actual structure.

The IFEFFIT XAFS data analysis suite [6] is used to simulate the EXAFS experiments. FEFF6 is used to simulate an XAFS experiment and IFEFFIT does post processing of the simulated EXAFS spectra. During the atomic structure refinement, the generated atomic structures will be run through these applications to obtain an EXAFS spectra.

### C. Known Models

Existing molecular structures that were generated using refined molecular quantum mechanics/mechanics chemistry (R-QM/MM) [1], and molecular quantum mechanics/mechanics chemistry and density functional theory (DFT-QM/MM) [1], will be used as a benchmark for comparison in this work. These are summarized in Section V.

## III. The evolutionary algorithm

### A. Algorithms

A collection of different evolutionary algorithms have been used in this work. The main focus of this study is the use of a Genetic Algorithm (GA) and a variation of Genetic Algorithm known as a Restarting GA (RGA). Two addition evolutionary algorithms, Particle Swarm Optimization (PSO) [7] and Differential Evolution (DE) [8] have been used as a post-optimization.

*1) Genetic Algorithm:* The GA has shown success in past studies in finding low-energy protein conformations. For example in [9] a GA is used to search through a continuous space in order to find low-energy protein conformations. This bears a similarity to this problem, as it is a continuous search space produces 3D structures. Also a population based search is important so that a number of candidate solutions are available, as some structures may be chemically unreasonable.

*2) Restarting Genetic Algorithm:* The RGA is a variation of the Recentering-Restarting Genetic Algorithm (RRGA) [10] [11] which has success in avoiding local minima. The RRGA is used to avoid fixating on local optima. RRGA does a series of GA runs which involves restarts and adjustments to the starting population. The RRGA selects a center, which is a possible candidate solution to the problem, and at the end of each basic GA run the center is compared to the best individual in the population. If the best individual is better than the current center it is replaced with the best individual and the whole process is repeated.

The RGA works similarly to the RRGA but there is no center for the population. Instead a basic GA is run until the population begins to converge. After a specified convergence percentage is reached, new individuals are added to the population. For example, if there are 100 individuals and the convergence rate is 5% then after all the duplicates are removed there will only be 5 individuals remaining. In restarting, shown in Algorithm 1, duplicate individuals were replaced with new individuals that have not yet entered the population. The implementation of restarting used in this work is described more fully in Section III-H.

**Algorithm 1** Restarting the population

```
if population has converged to minimum diversity then
    remove all duplicate individuals;
    while population not full do
        insert random draw from generated individuals into
        population;
    end while
end if
```

TABLE I. Sample Chromosome Representation

| X | Y | Z |
|---|---|---|
| 14.451 | -13.346 | 1.133 |
| 15.336 | -13.488 | 2.014 |
| 13.005 | -13.364 | 1.452 |
| 0.019 | 0.011 | 0.045 |
| ... | ... | ... |

### B. Representation

*1) Genetic Algorithm:* The chromosome representation consists of a list of 3-dimensional coordinates in space, where each position is assigned to a specific atom in the atomic structure. The actual position of each atom is not relevant because the goal is to determine the relative distances between the atoms. A subset of an individual can be seen in Table I. The units of measurement for each atom position are measured in Angstroms (Å).

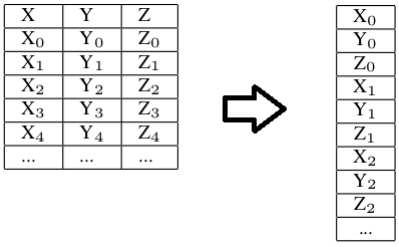
*2) Particle Swarm Optimization and Differential Evolution:* The individuals for the PSO and DE had to be modified to better suit these algorithms. Figure 1 demonstrates how the individuals were converted from a list of 3-dimensional positions to a single list of floating point values. In order to evaluate the fitness the list was translated back to a list of 3-dimensional positions.

### C. Fitness

The goal of the experiment is to find an atomic structure that generates the same EXAFS spectra as the experimental EXAFS spectra. If an atomic structure were to be generated that produces the same EXAFS spectra as the one found in experimentation then it can be said that the real atomic structure would have been found.

To calculate how close the calculated EXAFS spectra is to the experimental EXAFS spectra, root-mean-square deviations

Fig. 1. Population Individual Modification



(RMSD), see Equation 1, will be computed between the calculated and experimental spectra. Each spectra is recorded at increments of 0.05 $k$ / $A^{-1}$ which allows $EXAFS\chi k^3$ to be compared at each increment. The goal is to get the RMSD value as low as possible. It is not reasonable to expect the RMSD to be zero, because the experimental EXAFS spectra isn't perfect. The environment in which the EXAFS spectra is recorded creates small errors in the result.

$$\sqrt{\frac{\sum_{t=1}^{n}\left(x_{1,t}-x_{2,t}\right)^2}{n}} \tag{1}$$

### D. Population

The initial OEC atomic structure came from the crystallographic PSII structure [12]. It is available in the Protein Data Bank (PDB) [13] as PDB ID 3ARC. It was used as a starting point for the experiments. To ensure that the atomic structure was as stable as possible, the structure was put into a molecular dynamics simulation and the atoms were allowed to move freely in space until the overall temperature of the system was reasonably low. This acted as the baseline atomic structure for all tests. NAMD [14] was used for the molecular dynamics simulations.

In order to generate the initial population, two different approaches were used. At first, the individual atom coordinates were randomly shifted to generate a population but during the EXAFS analysis too many invalid atomic structures were being generated. The second approach taken was to let a molecular dynamics simulation generate an initial population. In this approach, the atomic structure was run through the molecular dynamics simulation again but instead the temperature of the system would be increased in order to make the atoms oscillate. The molecular dynamics simulation was allowed to run for 10000 steps. Each of these steps represents a unique atomic structure of the OEC and these were used to seed the genetic algorithm's population.

The atomic structures that were generated contained 1269 chemical elements. For the purposes of OEC structure refinement only 79 specific atoms were required for EXAFS analysis. The genetic algorithm only used the 79 atoms that required refinement. The generated atomic structures were run through IFEFFIT and compared to the target EXAFS spectra. The top 3% (roughly 300) were used to generate the initial populations in the genetic algorithm.

### E. Crossover operator

Two individuals are selected from the population and one-point crossover is performed. A random integer value is chosen based on the number of atoms in the individual, and the two individuals will trade their atomic positions based on the pivot point. One point crossover was chosen because it is less destructive to the individuals than other forms of crossover.

### F. Mutation operator

For the mutation operator a single atom coordinate will be moved. A random atomic coordinate is selected from the individual and its position is moved randomly by 0.05Å using Euclidean distance. The resulting position will be 0.05Å away from its original position. In order to determine how much distance the atomic position should be moved, an analysis was needed to learn more about how changing atomic positions affects the calculated EXAFS spectra.

TABLE II. MINIMUM MOVE REQUIRED AT 1%

| Element | 1% Difference | 5% Difference |
|---|---|---|
| O | 0.025Å | 0.5Å |
| Mn | 0.01Å | 0.5Å |
| Ca | 1Å | 5Å |
| C | 0.5Å | 5Å |
| N | 0.5Å | 5Å |
| H | 5Å | 5Å |

The analysis consisted of moving each atom, individually, in a variety of directions and calculating its RMSD score. Each atom was moved in a total of six directions ($\pm$X, $\pm$Y, and $\pm$Z), at a variety of distances (0.001Å, 0.005Å, 0.01Å, 0.025Å, 0.05Å, 0.1Å, 0.5Å, 1Å, and 5Å). This was done to determine how much movement was required of an atom to make a significant change to the RMSD score. Table II shows results of how much movement is required to produce a 1% and 5% change to their RMSD scores. Since there is more than one instance of each chemical element in OEC, the distance chosen was the first distance that produced the minimum change because the goal was to find the absolute minimum for each chemical element.

The value of 0.05Å was chosen for the experiments as a middle ground that could be applied to each chemical element. It should be noted that the value of 0.05Å is particular to OEC. A similar analysis could be done to determine the minimum move distance for each element in another chemical complex.

### G. Selection operator

For the selection operator a 3-tournament selection was used.

### H. Restarting

For the implementation of RGA, each run was seeded with 300 unique individuals. The initial population only used the specific amount of individuals it needed at generation 0 and during each population restart duplicates in the population re replaced with new individuals from the original 300. Figure 2 demonstrates how a population looks while restarting is occurring. Restarting has occurred at generations 19, 35, 46, and 71. The periodic decreases in fitness are destructive to the population but overall a consistent improvement in the fitness is seen throughout the run. The new individuals that are injected into the population allow for continued improvement to the fitness.

### I. Experiments

The GA and RGA will be performed as direct comparisons with each other. The variety of system parameters tested can be seen in Table III. A population size of 50 was selected as a result of empirical testing. The number of runs was decided to be 10 in order to obtain sensible results in a reasonable amount of time.

Fig. 2. Results from a sample run using the population restarting technique

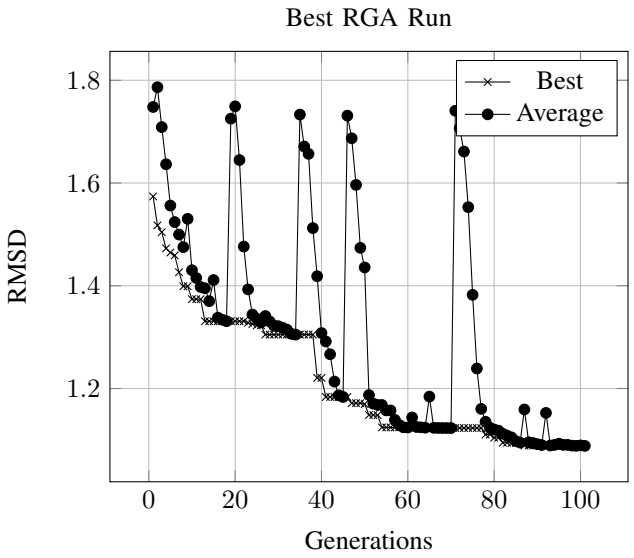



TABLE III. SYSTEM PARAMETERS USED IN GA AND RGA

| Settings | Values |
|---|---|
| Runs | 10 |
| Generations | Max. 30, Until Converged |
| Population Size | 50 |
| Crossover Rate | 80, 70 |
| Mutation rate | 10, 20, 30 |
| Elitism | True, False |
| Number of restart attempts | 3, 5 |
| Max convergence percentage before restarting | 5%, 10% |

### J. Results and Discussion

Table IV contains the best results of the GA and RGA experiments. Immediately it was clear that the RGA was outperforming the GA. A closer look revealed that the basic GA experiments were converging early on local optima. The RGA experiments initially converged on similar optima but were able to find new optima after each restarting phase. The RGA performed better than the GA but required more than double the amount of generations. Figure 2 provides an example of the generational data for the best run of RGA.

The average RMSD score from the RGA was able to outperform both benchmark RMSD scores but only by a small margin. Fortunately, because the goal is to find a single individual, the best candidate solution from the RGA was able to far exceed the benchmark scores with RMSD of 1.0877. Figure 3 shows the comparison of the EXAFS spectra.

After examining the calculated EXAFS spectra and the experimental EXAFS spectra it becomes clear that the calculated EXAFS spectra will never exactly fit the experimental EXAFS spectra. In particular, there appears to be many jagged jumps along the experimental EXAFS spectra that create many difficulties for matching by the calculated EXAFS spectra. Nonetheless, this can produce a good approximation for further experimentation.

TABLE IV. RESULTS FROM GA AND RGA EXPERIMENTS

| Gen. | Crossover | Mutation | Elitism | Conv. Rate | Restarting | Avg. Best |
|---|---|---|---|---|---|---|
| 30 | 80% | 20% | False | - | - | 1.3411 |
| 30 | 80% | 10% | False | - | - | 1.3746 |
| 30 | 70% | 30% | False | - | - | 1.2756 |
| 30 | 80% | 10% | True | - | - | 1.3278 |
| 30 | 70% | 30% | True | - | - | 1.2591 |
| 30 | 80% | 20% | True | - | - | 1.3299 |
| 67 | 80% | 20% | True | 10% | 3 | 1.2247 |
| 81 | 80% | 20% | True | 5% | 3 | 1.2117 |
| 107 | 80% | 20% | True | 10% | 5 | 1.2062 |
| 124 | 80% | 20% | True | 5% | 5 | 1.2156 |
| 67 | 70% | 30% | True | 5% | 5 | **1.1998** |
| 81 | 70% | 30% | True | 10% | 5 | 1.2138 |

TABLE V. LOCAL SEARCH RESULTS

| Algorithm | Initial Movement Radius | Pop. Size | Gen. | Average Best |
|---|---|---|---|---|
| PSO | $\pm$0.05Å | 50 | 100 | 0.7782 |
| PSO | $\pm$0.05Å | 50 | 30 | 0.8976 |
| PSO | $\pm$0.25Å | 50 | 30 | 1.2411 |
| DE | $\pm$0.05Å | 50 | 30 | 1.1354 |
| DE | $\pm$0.25Å | 50 | 30 | 1.7220 |

### K. Post-Optimization

Post-optimization was performed because the search space is too large to search every possible configuration. Algorithms based on continuous space problems were not very successful at developing good solutions to the problem. PSO and DE were originally tested against the RGA performance but the results were not promising. Once the RGA was able to produce high quality solutions for OEC these candidate solutions were applied to PSO and DE as an attempt at post-optimization.

A candidate solution found from the GA experiments was used to seed the initial population of the PSO and DE algorithms. The peer-reviewed computational intelligence library, CILIB [15], was used to run the PSO and DE experiments.

To generate the initial population, for both PSO and DE, each floating point value within each individual was randomly adjusted by a range of numbers. At first adjusting each value by $\pm$1Å was tried but the results were far worse than the original candidate solution. Improvements did not begin to appear until the adjusted amount was significantly reduced. Ranges of $\pm$0.05Å and $\pm$0.25Å were used to generate improved results. Table V shows the results of the two post-optimization strategies. These scores are averaged over 30 runs.

The results of the DE experiments were not promising. Results were improving but their RMSD scores never improved past the candidate solutions. In contrast, the results of the PSO experiments were extremely successful. Figure 4 shows the generational results of the PSO averaged over 30 runs. The RMSD score of the candidate solution was significantly reduced which shows that PSO works very well as a post-optimization strategy.

Fig. 3. OEC EXAFS Spectra Comparison

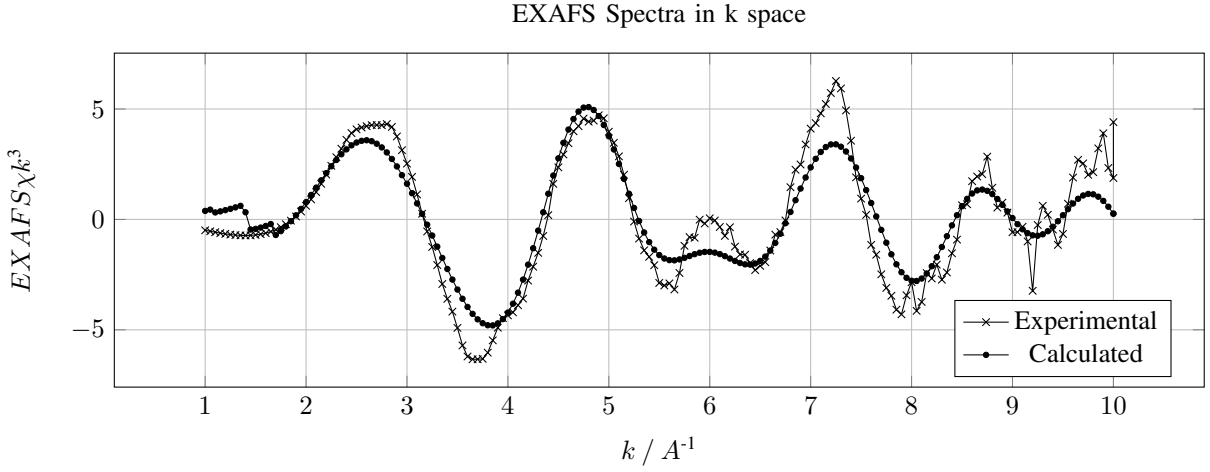
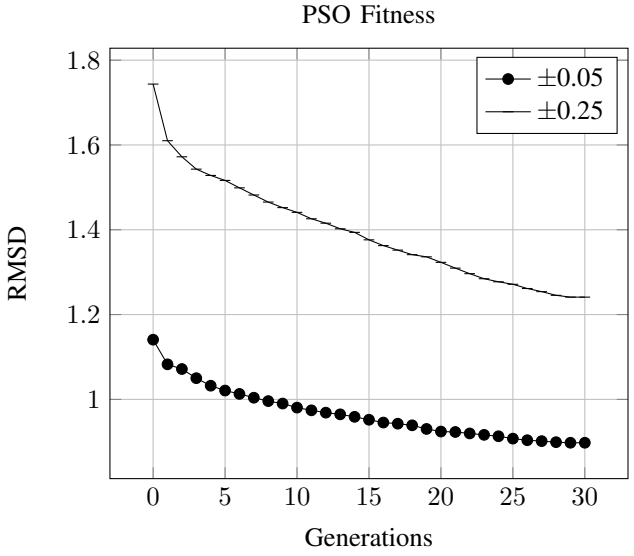



Fig. 4. PSO Post-Optimization Results




The results from the PSO and DE experiments using different initial movement radii were very surprising. Using a smaller initial movement radius suggests that the initial candidate solution is already very close to the ideal solution. The PSO was able to improve the best candidate by almost 50%. Figure 5 shows the best EXAFS spectra after optimization from the PSO which has an RMSD score of 0.7287.

## IV. Subset Testing

In an attempt to reduce the search space of the problem, different chemical elements were left rigid during the evolutionary process. Table VI gives a breakdown of the atoms contained within OEC.

Experiments were run on different subsets of chemical elements. Each subset was allowed to moved freely while the chemical elements not listed were kept rigid. Table VII demonstrates the results over 10 runs. Certain rows in Table VII contain **N/A** as their values were too large. The atoms listed in each row were allowed to be flexible during the experiment. The parameters were kept consistent for each experiment and can be seen in Table VIII.

TABLE VI. Chemical Element Breakdown

| Element | Count |
|---|---|
| Mn | 4 |
| Ca | 1 |
| O | 26 |
| C | 14 |
| N | 6 |
| H | 28 |
| Total | 79 |

TABLE VII. Experiments with different subsets

| Atoms | Best | Average |
|---|---|---|
| Mn, Ca, C, O, N, H | 1.1998 | 1.2580 |
| Mn, Ca, C, O, N | 1.1697 | 1.2621 |
| Mn, Ca, C | 2.4413 | N/A |
| Mn, Ca, O | 1.2531 | N/A |
| Mn, Ca, N | 2.4881 | N/A |
| Mn, Ca, C, O | 1.1648 | N/A |
| Mn, Ca, C, N | 2.5143 | 2.5649 |
| Mn, Ca, O, N | N/A | N/A |
| Mn, Ca | 2.4916 | 2.5088 |

The results of the experiments revealed some interesting characteristics. Having the Hydrogen elements rigid actually produces improved results. This is not surprising since early examination of the atoms revealed that moving a Hydrogen

TABLE VIII. GA Subset Parameters

| | |
|---|---|
| Runs | 10 |
| Population size | 50 |
| Crossover rate | 0.7 |
| Mutation rate | 0.3 |
| Elitism | True |
| Number of restart attempts | 5 |
| Max convergence percentage before restarting | 5% |

Fig. 5. OEC EXAFS Spectra Comparison for PSO Post-Optimization

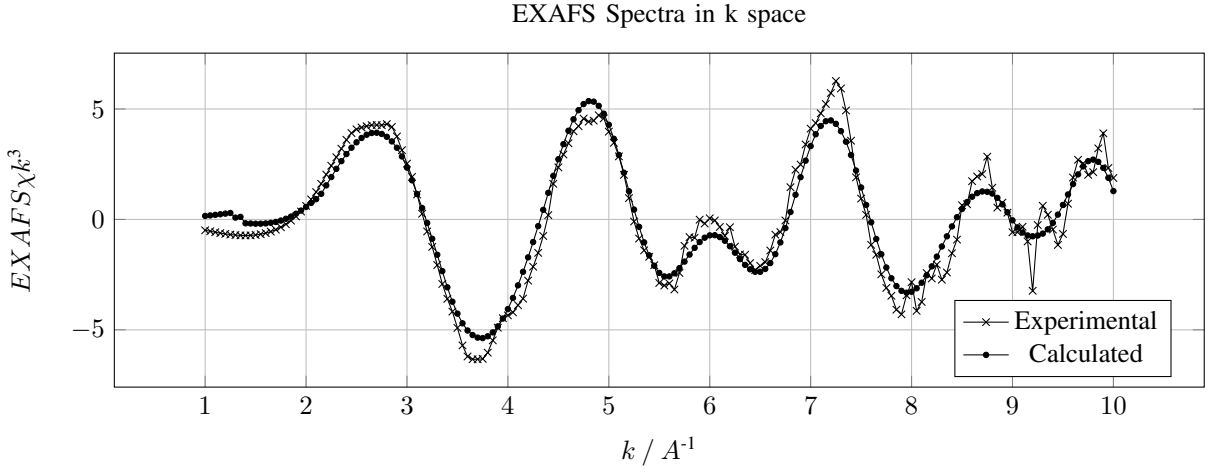



TABLE IX. COMPARISON OF RESULTS

| Algorithm | Best Results |
|---|---|
| DFT-QM/MM [1] | 1.2643 |
| R-QM/MM [1] | 1.2403 |
| RGA | 1.0877 |
| PSO | 0.7287 |

element had very little impact on the RMSD score. It is suspected that the RMSD score has improved because the chromosome length has been reduced from 79 to 51. The reduced chromosome length would allow for more different combinations to be attempted. This idea will be examined in more detail in future work, as discussed in Section V.

## V. CONCLUSION AND FUTURE WORK

This work shows that the GA and RGA were able to produce comparable atomic structures to those of existing solutions. The RGA was able to outperform existing solutions but only by a small margin. Using PSO as a post-optimization technique was able to significantly improve the solutions found by the RGA. Table IX shows a comparison of the best results found using RGA and PSO.

In the initial stages of this work PSO and GA were used for testing but the PSO was not able to perform with any reasonable success. With the improvements the PSO was able to make to the RGA results, more testing is required using the PSO. It is possible that the initial poor results from the PSO are due to the initial population. In the future an alternative method of generating a population for PSO and DE should be considered.

The study performed in Section III-F was done to determine how much an atom should be shifted during the mutation phase. The analysis revealed that different chemical elements require different minimum distances to create a noticeable effect on the RMSD score. In this work the each atom is treated the same way and shifted by the absolute minimum distance found. Future work should attempt to shift each atom by their own minimum distances, thereby allowing for more precise results.

The molecular dynamics simulation performed to create the initial population was useful in creating candidate solutions for the GA. Using the molecular dynamics simulation should be studied as a way of restarting the GA. Once the best candidate solution is found using the GA it should be placed back in the molecular dynamics simulation to generate new atomic structures that could be used to restart the GA.

It should be noted that the results in this work have not been tested for chemical feasibility. In particular, subset testing was performed as a proof of concept for future work. The exploration of subset testing in this work may be producing results that are chemically infeasible. To solve this problem, the created atomic structure should be place in the molecular dynamics program, discussed in Section III-D, to optimize the positions of the rigid chemical elements.

A secondary goal, that is not explored in this work, of atomic structure refinement is to generate atomic structures that produce a low potential energy. Some configurations of the atomic structure produce chemically unreasonable structures meaning that their potential energy is too high or they cannot exist in nature. The atomic structures should be analyzed using force fields to measure their feasibility and potential energy.

A benefit of using a population based algorithm is that at the end of a run there are several possible candidate solutions. For the purposes of atomic structure refinement using EXAFS Spectra analysis having a set of candidate solutions is actually beneficial because each candidate solution is going to have different chemical properties. Further analysis of these individuals, by an expert, could expose which candidates are most suitable.

A multi-objective approach could be applied to molecular structure refinement since the ideal model would have both a low RMSD score and low potential energy. The cost of adding the force field calculations would double the time required to evaluate an individuals fitness, which is already significantly high. Multi-objective experiments may only require a models potential energy to be calculated at certain intervals to reduce the evaluation time but this still demands research.

## ACKNOWLEDGMENT

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada.

The authors would like to thank Christian Negre for providing EXAFS Spectra data for OEC $S_1$ and James Hughes for helpful comments on preparing this work.

## REFERENCES

[1] S. Luber, I. Rivalta, Y. Umena, K. Kawakami, J.-R. Shen, N. Kamiya, G. W. Brudvig, and V. S. Batista, "S1-state model of the o2-evolving complex of photosystem ii," *Biochemistry*, vol. 50, no. 29, pp. 6308–6311, 2011.

[2] Y.-W. Hsiao, Y. Tao, J. E. Shokes, R. A. Scott, and U. Ryde, "Exafs structure refinement supplemented by computational chemistry," *Physical Review B*, vol. 74, no. 21, p. 214101, 2006.

[3] E. M. Sproviero, J. A. Gascón, J. P. McEvoy, G. W. Brudvig, and V. S. Batista, "A model of the oxygen-evolving center of photosystem ii predicted by structural refinement based on exafs simulations," *Journal of the American Chemical Society*, vol. 130, no. 21, pp. 6728–6730, 2008.

[4] W. Banzhaf, P. Nordin, R. E. Keller, and F. D. Francone, "Genetic programming: An introduction: On the automatic evolution of computer programs and its applications (the morgan kaufmann series in artificial intelligence)," 1997.

[5] M. Newville, "Fundamentals of xafs," *Consortium for Advanced Radiation Sources, University of Chicago (USA)[http://xafs. org]*, 2004.

[6] Ifeffit: Interactive xafs analysis. [Online]. Available: http://cars9.uchicago.edu/ifeffit/Ifeffit

[7] J. Kennedy, "Particle swarm optimization," in *Encyclopedia of Machine Learning*. Springer, 2010, pp. 760–766.

[8] R. Storn and K. Price, "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces," *Journal of global optimization*, vol. 11, no. 4, pp. 341–359, 1997.

[9] P. Comte, "Bio-inspired optimization & sampling technique for side-chain packing in mcce," Mastersthesis, Brock University, 2010.

[10] J. Hughes, S. Houghten, and D. Ashlock, "Recentering, reanchoring & restarting an evolutionary algorithm," in *Nature and Biologically Inspired Computing (NaBIC), 2013 World Congress on*. IEEE, 2013, pp. 76–83.

[11] J. Hughes, J. A. Brown, S. Houghten, and D. Ashlock, "Edit metric decoding: Representation strikes back," in *Evolutionary Computation (CEC), 2013 IEEE Congress on*. IEEE, 2013, pp. 229–236.

[12] Y. Umena, K. Kawakami, J.-R. Shen, and N. Kamiya, "Crystal structure of oxygen-evolving photosystem ii at a resolution of 1.9 å," *Nature*, vol. 473, no. 7345, pp. 55–60, 2011.

[13] Crystal structure of oxygen-evolving photosystem ii at 1.9 angstrom resolution. PDB Protein Data Bank. [Online]. Available: http://www.rcsb.org/pdb/explore/explore.do?structureId=3arc

[14] Namd - scalable molecular dynamics. Theoretical and Computational Biophysics Group. [Online]. Available: http://www.ks.uiuc.edu/Research/namd/

[15] C. I. R. Group. Computational intelligence library. [Online]. Available: http://www.cilib.net/