# Investigating Large-Amplitude Protein Loop Motions as Extreme Events Using Recurrence Interval Analysis

**البحث في حركات حلقة البروتين ذات السعة الكبيرة باعتبارها أحداث قصوى باستخدام تحليل فترات التكرار**

**By:**

**Yasmeen Ali Ashour**

**Thesis committee:**

**Prof. Wael Karain (Principal advisor)**

**Dr. Hazem Abu Sara (Member)**

**Dr. Abdallah Sayyed Ahmad (Member)**

**This thesis was submitted in partial fulfillment of the requirements for the Master's degree in Physics from the Faculty of Graduate Studies at Birzeit University, Palestine**

**27/5/2020**

# Investigating large-amplitude protein loop motions as extreme events using recurrence interval analysis

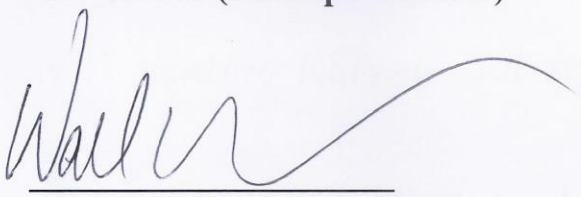
**A thesis submitted in partial fulfillment of the requirements for the Master's degree in Physics from the Faculty of Graduate Studies at Birzeit University, Palestine**

**By:**

**Yasmeen Ali Ashour**

**Thesis committee:**

**Prof. Wael Karain (Principle advisor)**



**Dr. Hazem Abu Sara**

**Dr. Abdallah Sayyed Ahmad**

**27/5/2020**

# *Dedication*

To a strong and gentle soul who taught me to trust in believe in hard work and that so much could be done with little, my mother. To who supporting encouraging to believe in myself and who was sacrificed all his energy to get where we are now, my father. To whom was the bond in my scientific journey and spared no effort in helping, dear husband. To my husband's family for their love, kindness and constant support. To my precious daughter and my first gladness my princess 'Jana'. To my wonderful sisters and brothers. To all friends.

# Acknowledgments

First of all, I would like to thank Almighty Allah for everything in my life. Without his guidance I would never be able to accomplish in my whole life.

I am also grateful to the supervisor, Prof. Wael, for his support and patience. I would like to thank members of my committee for reviewing my thesis. Thank you for your helpful feedback.

# Table of content

## Table of figures

## Abstract

We investigate large amplitude motions of 106 loop pairs of residues in β-lactamase inhibitor protein BLIP as extreme events using recurrence interval analysis. We prepare probability distribution functions $P_q(\tau)$ of recurrence intervals $\tau$ between two consecutive time steps in the distance time series above a range of positive threshold $q > 0$. We fit these distribution using power laws, and calculate the significance values for the power model parameters. After that, we apply hazard probability functions $W(t|\Delta t)$ to forecast the occurrence of large volatility events. Here are the main results we were obtained:

i. The power law exponents lie in the range [1.46 – 2.42] with most being in the range [1.70 – 1.79].
ii. The probability distribution for the 106 pairs of residues are power law fitted with the $p-$values ranging from 0 to 0.81. Many p-values exceed 10% significance level.
iii. Using a " receiver operator characteristic " (ROC) analysis, we predict the probability of extreme event occurrence. Most of the prediction rates (97%) exceed the random prediction rate.

## ملخص

نحن نبحث في حركات ذات الاتساع الكبيرة لـ 106 أزواج حلقة من الأحماض الأمينية في بروتين مثبط $\beta$-lactamase باعتبارها أحداث قصوى باستخدام تحليل فترات التكرار. نقوم في هذا العمل التوزيعات الاحتمالية $P_q(\tau)$ لفترات التكرار $\tau$ بين خطوتين زمنيتين متتاليتين في سلسلة المسافة الزمنية فوق جهد قيمة عتبة موجبة $q> 0$. ثم نلائم هذه التوزيعات باستخدام قوانين الطاقة الأسية ونحسب القيم المعنوية لمعاملات نموذج قانون الطاقة. بعد ذلك ، نطبق اقترانات احتمالية الخطر $W(t \mid \Delta t)$ للتنبؤ بحدوث أحداث ذات تقلبات كبيرة. و فيما يلي النتائج الرئيسية التي حصلنا عليها:

1- قيم الأسس لقانون الطاقة تقع في الفترة [1.46-2.42] بينما معظمها تركز في الفترة [1.70-1.79].

2- التوزيع الاحتمالي لـ106 أزواج من الأحماض الأمينية هي عبارة عن قانون طاقة يتناسب بقيم $p - value$s تتراوح من 0 إلى 0.81 . العديد من قيم $p$ تتجاوز القيمة المعنوية 10%.

3- باستخدام تحليل "خصائص عامل الاستقبال" (ROC) ، نتوقع احتمالية وقوع الحدث الحرج . معظم معدلات التنبؤ (97٪) تتجاوز معدل التنبؤ العشوائي.

# 1. Introduction and Background

## 1.1 Protein Loop

Proteins are made up of three types of secondary structure elements: α-helices, β–sheets (parallel or antiparallel), and loops. Fig 1.1 shows a section of a protein loop, which serves as a connecter between α-helices and β–sheets [1].

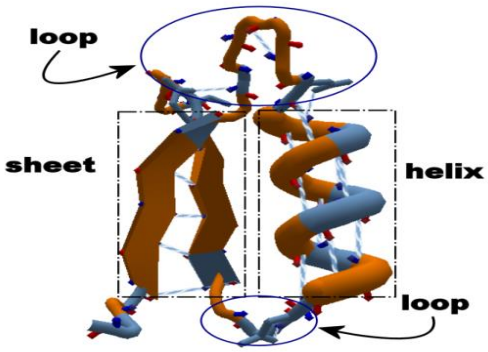



**Fig 1. 1:** A protein loop (in blue circle). This loop connects an α-helix and a β-sheet (in dashed boxes)[2].

Loops lie on the surface of a protein [1] and are thus exposed to the surrounding solvent [3]. The nature of a protein loop is variable unlike α-helices and β–sheets. Loops are classified as a short if they have ten or less amino acids , and long if they have more than 10 residues [4]. Loop flexibility plays an important role in protein function[5]. Loops also work to shelter the active site residues [6], and to transport ions or molecules from secondary binding sites towards their primary binding site[1].

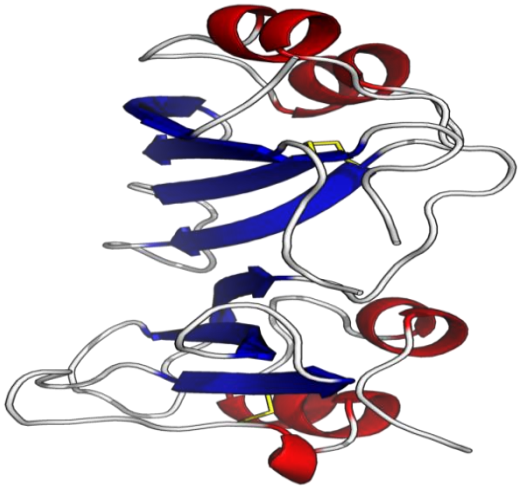


**Fig 1. 2:** Representation of Beta-Lactamase Inhibitory Protein (BLIP) with α-helices in red, β-sheets in blue[7].

In this work, we will study the Beta Lactamase Inhibitor Protein (BLIP) (Fig 1.2). BLIP is produced by a type of bacteria called *Streptomyces clavuligerus* and has 165 amino acid residues[8,9]. The flexibility domains allows it to inhibit β-lactamases such as TEM-1 , resist penicillin antibiotics. It also inhibits many of class A β-lactamases [10]. We will concentrate on four loops. The first loop L1 lies between the sheets β6 and β7 (residues 133-145). The second loop L2 lies between the sheets β2 and β3 (residues 48-49). The third loop L3 lies between the sheets β5 and β6 (residues 116-125). The fourth loop L4 lies between the sheets β3 and β4 (residues 59-66). In this work the group consisting of **L1** and **L2**, and **L3** and **L4**, will be referred to **G1** and **G2** respectively.

## 1.2 Recurrence interval between extreme events

The recurrence interval between extreme events can be defined as the time distance between two consecutive extreme events, either above a positive defined threshold $(+q)$, or below a negative defined threshold $(-q)$ [11]. Figure 1.3 shows six recurrence intervals. Three of them $(\tau_1, \tau_2, \tau_3)$ are above a threshold value of (+2). The other three $(\tau_4, \tau_5, \tau_6)$ are below a threshold value of (-2). The mathematical expression for these intervals in general for a positive threshold $q > 0$ is given by:

$$\tau(t) = \min\{t' - t : m(t') > q, t' > t\}. \quad (1.1)$$

Where $m(t')$ is the normalized time distance that exceeds the threshold q.

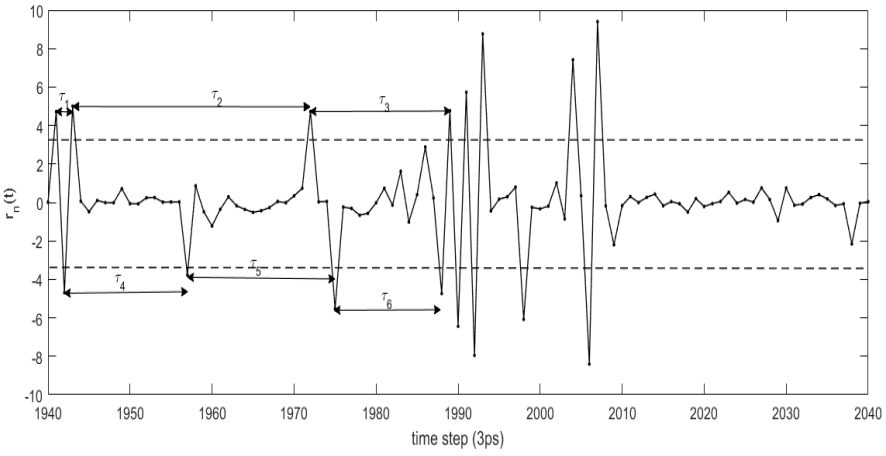



**Fig 1. 3:** Schematic illustration of six recurrence intervals for thresholds larger or lower than a given threshold **q**. In this example, thresholds are chosen to be $\mathbf{q > +2}$ **and** $\mathbf{q < -2}$.

The statistics of these recurrence intervals has attracted wide attention recently [12,13]. Applications span trading volumes [14], climate records [15,16], heartbeat intervals in medical science[17], financial volatilities [18,19], precipitation and river runoff [20]. If the dynamics of these intervals can be modeled, then one can use these models to predict the probability of occurrence for these extreme events.

Recently, Recurrence Interval Analysis RIA was used to study large amplitude motions of two residues in BLIP: ASP49 and PHE142 [11]. In this work we will use the same technique to study a larger group of residues falling within loops in BLIP. Specifically, we will study the two groups G1 and G2 that were previously introduced in section 1.1. These two groups consist of 106 different pairs of residues. The distance time series between the residues in each pair will be investigated. Fig.4 shows the distance between the center of masses of residues 59 and 116 in BLIP protein. This distance undergoes large fluctuation [11]. This time series is 300ns long, and consists of 100000 points, with a 3ps time interval between consecutive points.

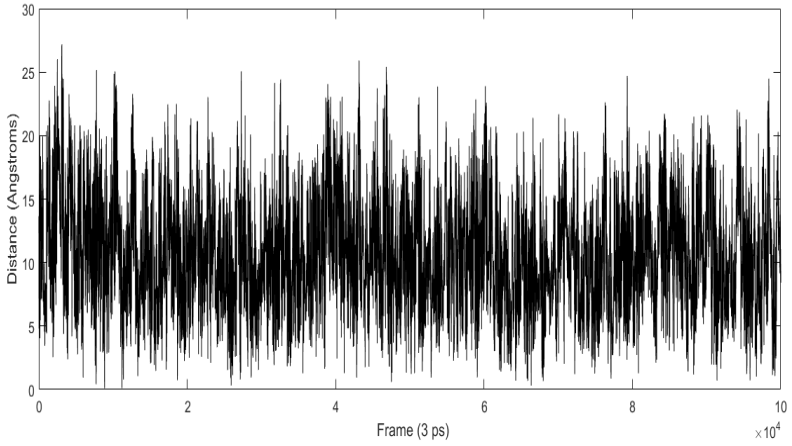



**Fig 1. 4:** The distance between the centers of mass of BLIP protein residues 59 and 116 over a 300 ns time period. Each time step is 3 ps long.

Fluctuations in distance are the deviation of the distances from their average values [21]. These fluctuations can be treated as extreme events that can be investigated using RIA.

We define the distance return as the logarithmic distance change between two consecutive time steps in the time distance series [22],

$$r(t) = ln\, d(t) - ln\, d(t-1) \tag{1.2}$$

where $d(t)$ is the distance between the centers of mass of two pairs of residues of $\beta$ −lactamase BLIP at time t.

The distance return is normalized by dividing by its standard deviation σ as follow [11]

$$r_n(t) = \frac{r(t)}{\sigma} \quad ; \quad \sigma = [< r(t)^2 > - < r(t) >^2]^{1/2} \tag{1.3}$$

Once this time series is prepared, we prepare arrays of all recurrence intervals $\tau's$ above six different defined positive thresholds: 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively. Then we will study the properties and the statistics of these recurrence intervals [22]. Fig. 1.5 shows a section of plots of the distance, logarithmic distance, and its normalized return.

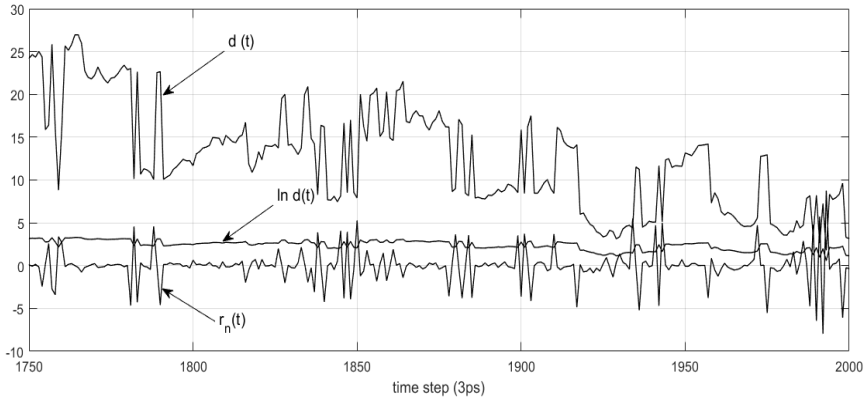



**Fig 1. 5:** Plots of the distance $\boldsymbol{d}(\boldsymbol{t})$, the logarithmic distance $\boldsymbol{lnd}(\boldsymbol{t})$ and it's normalize difference return $\boldsymbol{r_n}(\boldsymbol{t})$. The section plot for the residue 49 and 145 in $\boldsymbol{\beta}$ −lactamase BLIP.

Basically, we will prepare 636 arrays of return intervals, six arrays for every one of the 106 residue pairs. Each array will then be analyzed using RIA.

## 1.3 Molecular dynamics

Molecular dynamics (MD) is used to model the motions of atoms in proteins using classical Newtonian mechanics. [23,24,25].

In our MD simulation, we used NAMD and VMD [26,27]. The BLIP protein

(PDB entry 3 gmu) was used [28]. The periodic boundary conditions were used in a $80A^{\circ} \times 80A^{\circ} \times 80A^{\circ}$ box. The protein was solvated using 15264 TIP3P water (0.15M/ NaCl), and neutralized using $20\ Cl^{-}$ ions and $22\ Na^{+}$ions [11]. The Particle-Mesh-Ewald method was used to perform the electrostatic calculations. A switching function was used for non-bonded interactions with a switch distance of 10 Å and a cutoff distance of 14 Å. The simulation was performed at 1.0 ATM pressure with an integration step of 2.0fs. Langevin dynamics were used to control temperature and pressure. The protein was minimized using the conjugate gradient method for 5000 time steps. Then it was gradually heated in small temperature steps of 10K starting at 100K, until it reached the final simulation temperature of 310K. The equilibration period was 5ns long. The full length of the simulation was 300ns. Finally, the time series for the distance between the centers of mass of pairs of residues was calculated for 106 pairs of residues in BLIP protein for residue groups G1 and G2[11].

## 2. Methods

### 2.1 Probability distribution function PDF of recurrence interval of distance returns

The first step in studying the dynamics of the recurrence intervals, is to prepare the probability distribution function (PDF) $Pq\ (\tau\ )$ of recurrence intervals for the normalized distance return $r_n(t)$. This done for six thresholds: 1.0, 1.2, 1.4, 1.6, 1.8, 2.0, and repeated for the 106 residues pairs in G1 and G2. This done using a simple MATLAB script as follows (for each threshold):

1) If the value of the array $r_n(t)$ is larger than the corresponding threshold, it is replaced by the integer 1. Otherwise, a value of zero is placed in its place.
2) The consecutive return times are calculated between the array elements that have 1 stored in them. The value of the return time τ is placed in the corresponding elements.
3) Once the array processing is finished, a histogram is prepared for the whole array. The population for different values placed in this histogram.
4) The average return time $\tau_{avg}$ is calculated.
5) The histogram is normalized to a total value of 1.

Fig 2.1 and Fig 2.2 show the normalized probability distribution functions (pdfs) $P_q(\tau)$ for the distance recurrence intervals at thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively for different two pairs of residues. In Fig. 2.1, the pdf

for the residue pair 49 and 145 is shown. At the low end of the τ spectrum between 1 and 10, one can notice that the probability decreases slightly with increasing thresholds. At higher values of τ, the probability of occurrence decreases. This is normal for rare extreme events.

In Fig. 2.2, the pdf for the residue pairs 62 and 119 is shown. This in a similar fashion to the one above. In fact, this is a general pattern for all the pdfs calculated for all the residue pairs.

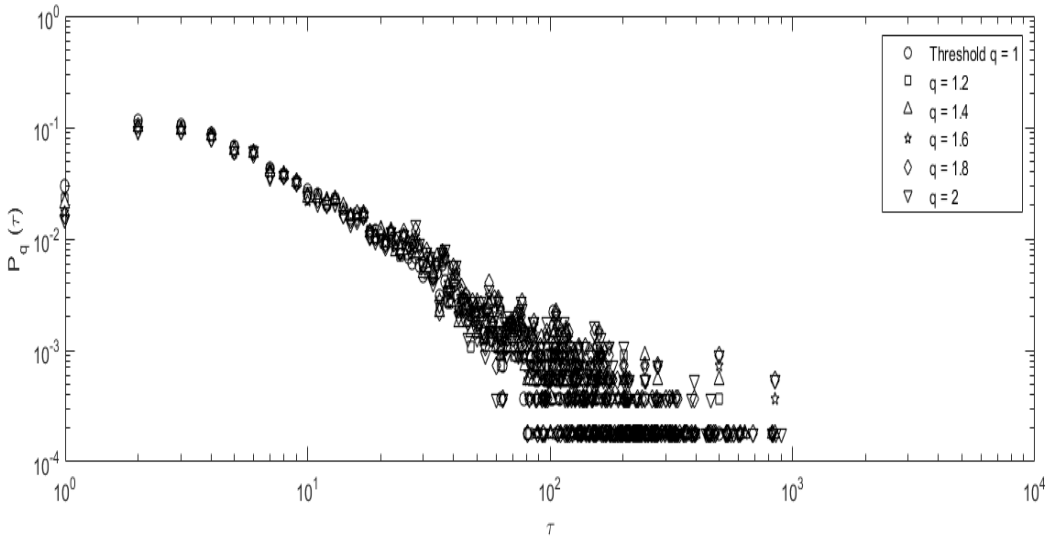



**Fig 2. 1:** Normalized probability distribution function of distance recurrence intervals between the centers of mass of residues 49 and 145. The distributions are calculated at the thresholds $\mathbf{q} = 1.0, 1.2, 1.4, 1.6, 1.8$, and 2 respectively.

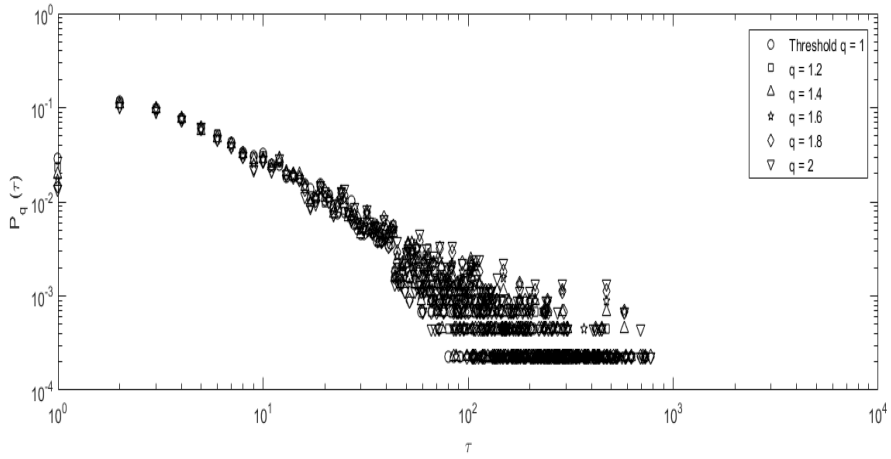



**Fig 2. 2:** Normalized probability distribution function of distance recurrence intervals between the centers of mass of residues 62 and 119. The distributions are calculated at the thresholds **q** = 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively.

The pdfs are scaled by multiplying them with the mean recurrence interval time $\tau_{avg}$[11,18]. In Fig. 8 and Fig. 9, two examples of scaled pdfs $P_q(\tau)\ \tau_{avg}$are plotted as a function of the scaled of recurrence interval $\tau/\tau_{avg}$ for two pairs of residues at thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively. We notice that scaled pdfs for different $q$ values collapse to a single curve, especially at small values of $\frac{\tau}{\tau_{avg}}$.

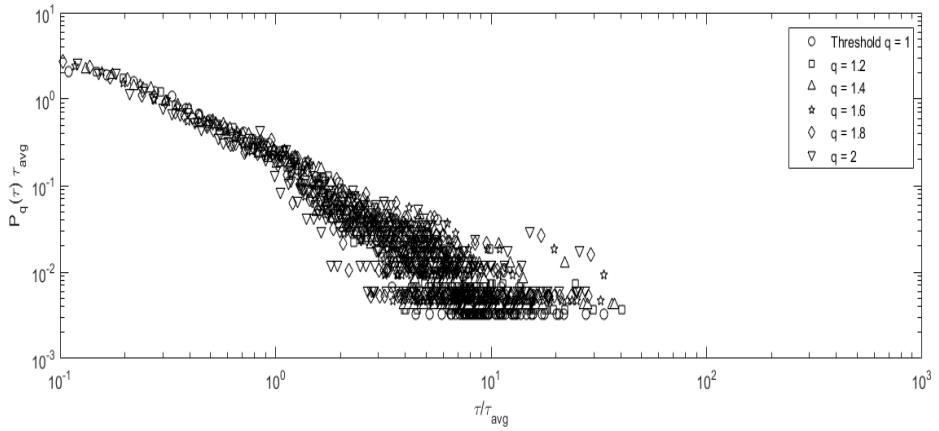



**Fig 2. 3:** Scaled pdfs $\boldsymbol{Pq(\tau)\ \tau_{avg}}$of distance recurrence intervals for the distance between the center of mass of residues 49 and 145. The distributions are calculated at thresholds $\mathbf{q}\ =$ 1, 1.2, 1.4, 1.6, 1.8, and 2 respectively.

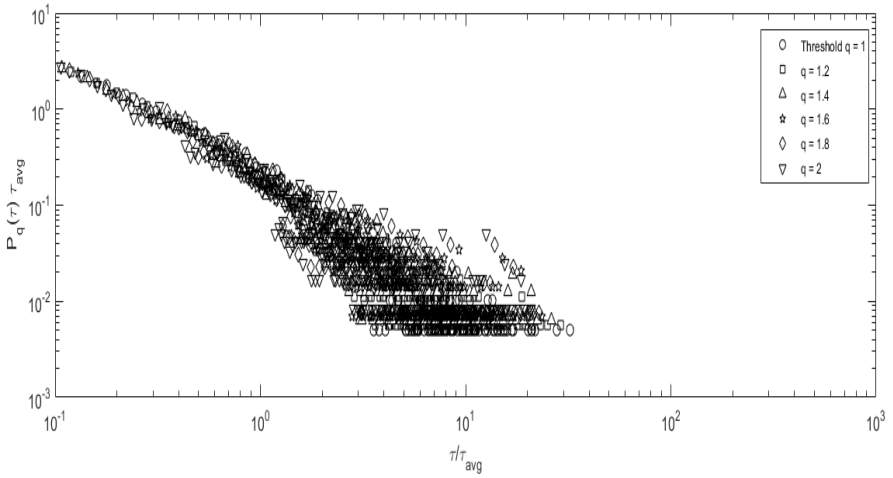



**Fig 2. 4:** Scaled pdfs $\mathbf{Pq(\tau)\ \tau_{avg}}$of distance recurrence intervals for the distance between the center of mass of residues 62 and 119. The distributions are calculated at thresholds $\mathbf{q}\ =$ 1, 1.2, 1.4, 1.6, 1.8, and 2 respectively.

## 2.2 Fitting the scaling function of recurrence interval PDFs

The linear behavior of the scaled pdfs in Fig 2.3 and Fig 2.4, on log-log plot, suggests that the scaled distribution follows a power law form. We can assume that the observed scaled pdfs above some $x_{min}$, which is the lower bound for the power law, follows [29].

$$f\left(\frac{\tau}{\tau_{avg}}\right) = f(x) = c\,x^{-\alpha},\ \ x \geq x_{min} \tag{2.1}$$

(The probability distribution diverges at zero, this means there should be a lower bound $x_{min} > 0$ for this distribution[30]). Three parameters need to be determined: the scaling parameter α. The normalization constant c, and the lower bound $x_{min}$.

Assuming that the scaling parameter $\alpha > 1$, the normalization constant is easily determined by integrating $\int_{x_{min}}^{\infty} cx^{-\alpha}dx = 1$, and $c = \frac{\alpha-1}{(x_{min})^{1-\alpha}}$ .

The scaling parameter is determined using the maximum likelihood estimator [31]. For example, if we have $N$ observed value $x_1 \ldots x_N \geq x_{min}$, the probability that these values come from a power law model is proportional to

$$p(x|\alpha) = \prod_{i=1}^{N} \frac{\alpha-1}{x_{min}} \left(\frac{x_i}{x_{min}}\right)^{-\alpha} \quad (2.2)$$

The probability in eqn (2.2) is called the likelihood of the given data. It is easier to work with the logarithm of the likelihood $L$. Then we take the derivative of $L$ with respect to α. The result is the scaling parameter that maximizes the likelihood function [30,32]

$$L = \ln p(x|\alpha) = ln \prod_{i=1}^{N} \frac{\alpha-1}{x_{min}} \left(\frac{x_i}{x_{min}}\right)^{-\alpha} \quad (2.3a)$$

$$= \sum_{i=1}^{N} [\ln(\alpha - 1) - \ln x_{min} - \alpha ln \frac{x_i}{x_{min}}] \quad (2.3b)$$

$$= Nln(\alpha - 1) - Nlnx_{min} - \alpha \sum_{i=1}^{N} ln \frac{x_i}{x_{min}} \quad (2.3c)$$

Finally, we solve $\frac{\partial L}{\partial \alpha} = 0$ for $\alpha$, then the result is:

$$\acute{\alpha} = 1 + n \left[ \sum_{i=1}^{n} \ln \frac{x_i}{x_{min}} \right]^{-1} \quad (2.4)$$

where $x_i$ are the observed value of $x$ in which $x_i \geq x_{min}$.

## 2.3 Estimating the lower bound on power-law behavior

The lower bound $x_{min}$ is the starting point of the power law behavior [30]. The method proposed by Clauset *et al.* was used to find the value of $x_{min}$[33]. The basic idea is to reduce the difference between the distribution of the observed data, and the best fit power- law model by using the Kolmogorov-Smirnov statistic K-S statistic [30] which is given by:

$$D = max_{x \geq x_{min}} |S(x) - P(x)| \quad (2.5)$$

Where $S(x)$ is the cumulative distribution function of the observed data and $P(x)$ is the cumulative distribution function of the estimated power law

distribution in the range $x \geq x_{min}$ .Then the value of $x_{min}$ is the value that minimizes $D$.

## 2.4 Goodness-of-fit tests

The final stage in any hypothesis is testing it so in this thesis we need to test how the estimated power law fits with empirical recurrence interval pdfs of 106 pairs of residues of BLIP. The test that was used is the goodness-of-fit using K-S [30,34]. For each scaled pdf, the KS value is calculated between the data and the power law model. A large set of synthetic datasets is then prepared using semi-parametric bootstrapping from the power law model. For each dataset, a power law model is then calculated. The KS distance is then calculated between each synthetic data set and its power law. The p-value is the fraction of times the resulting KS value is larger than that between the original data and its power law. The closer the p-value is to 1, the more confident one is that a power law does indeed describe the scaled pdf.

## 2.5 Hazard function and predictability

An important reason to study the extreme events using recurrence interval analysis is that it allows one to predict future extreme events with some confidence. For this goal, we define the " hazard function or hazard probability ". The hazard function is defined as the probability there will be another extreme during next time interval $\Delta t$ above defined threshold, where $t$ is the time of last previous extreme event that occurred above $q$ [35,36,37]. The

relation between the hazard function $W_q(t;\Delta t)$ and the probability distribution of the recurrence intervals between extreme events $P_q(t)$ is given by

$$W_q(t;\Delta t) = \frac{\int_t^{t+\Delta t} P_q(t)dt}{\int_t^{\infty} P_q(t)dt} \tag{2.6}$$

This is easily evaluated once the scaled pdfs $P_q(t)$ is defined as a power law. Hazard functions can be applied in many fields such as earthquakes, floods, stock returns, sales, and to forecast of coming extreme events [38]. After defining $\Delta t$ and threshold $q$, we use a decision-making algorithm to forecast extreme events [37]. An alarm will be set if the hazard probability exceeds a certain threshold [38]. To estimate the alarms performance, we should estimate the correct prediction rate and false alarm rate [28]. This can be achieved using the "receiver operating characteristic" curve(ROC). This is a plot of a correct prediction rate D versus a false alarm rate A [39,40]. To get D and A at a certain hazard probability threshold Q, we generate alarms and non-alarms at each time point [37]. We then compare these forecasting signals to the real data. We should get one of the following four cases: (1) a correct prediction of an extreme event, (2) a correct prediction of a non-extreme event, (3) a missed extreme event, (4) a false alarm. We define the following parameters: (i) number of extreme events correctly predicted $n_{11}$, (ii) number of missed events $n_{01}$, (iii) number of false alarms $n_{10}$, and (iv) number of non-extreme events correctly predicted $n_{00}$. After that, the correct predication rate $D$ and false alarm rate $A$ can be evaluated using the following equations:

$$D = \frac{n_{11}}{n_{01}+n_{11}} \tag{2.7a}$$

$$A = \frac{n_{10}}{n_{00}+n_{10}} \tag{2.7b}$$

By varying Q between 0 and 1, we find the pairs $(A, D)$ for each time series at the corresponding q thresholds. If $Q = 0$, then $A = D = 1$. If Q =1, then $D = A = 0$. The ROC curve connects point $(0,0)$ at the lower left corner, and the point $(1,1)$ at the top right corner. For random guesses, $D = A$, and a straight line is observed between these two points. The area under the ROC curve represented by $AUC$ is a measurement of predicting performance and can be calculated as follows [37]

$$AUC = \int_0^{0.3} D(A)dA \tag{2.8}$$

We stop the integration at 0.3 because higher false alarm rates are useless.

Thus our goal in this thesis is to model large changes in inter-residue distances as extreme events following a power law model. These models will then be used to calculate hazard functions. The hazard functions will then be used to predict extreme events in the distance time series. Their prediction capability will be analyzed and tested.

## 3. Results and discussion

Fig 3.1 shows the distribution of α-values for 106 pairs of residues at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 ( a total of 736 values). The range of α- values is [1.46 – 2.42] and a large portion of these fall between 1.51 and 1.92. (See Appendix A). In general, the α values increase with the threshold. We notice that the largest value of α is 2.42 at threshold 1.6 for residue pair 119-62 which falls in group G2.

Fig 3.2 shows the relationship between the scaling parameter α and the lower bound value $x_{min}$ for the 736 cases. In general, $x_{min}$ increase with α. The maximum value of $x_{min}$ (0.5373) is associated with the maximum value of α (2.42). We notice at thresholds 1.0, 1.2, and 1.4, that the value of $x_{min}$is small compared with $x_{min}$ at threshold 1.6, 1.8, and 2.

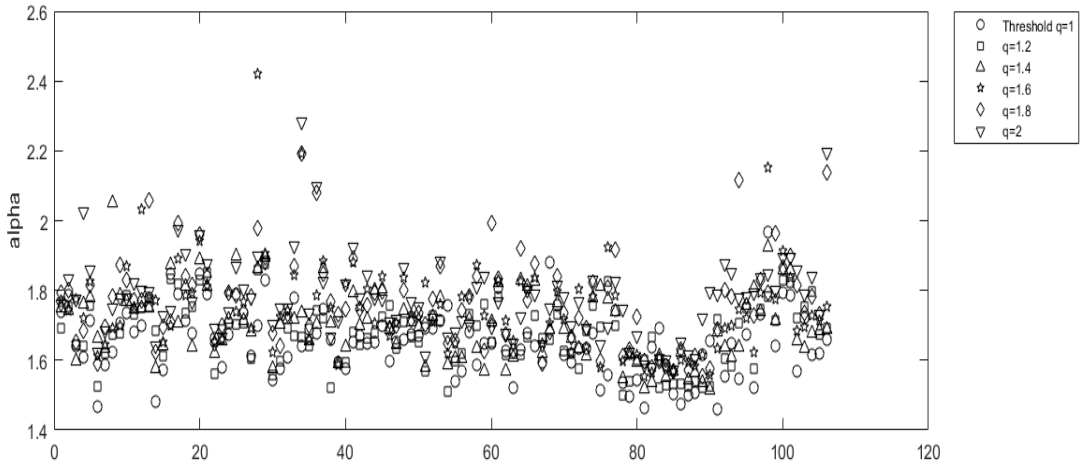



**Fig 3. 1:** The distribution α-values for 106 pairs of residues at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively.

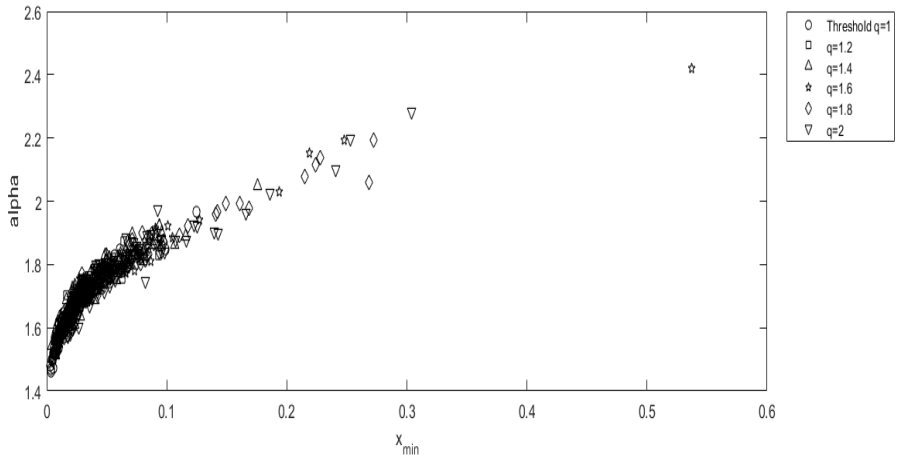



**Fig 3. 2:** The dependence between α and $\boldsymbol{x_{min}}$.

Fig 3.3 shows the distribution of p-values for 106 pairs of residues at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively. We notice that there are many p-values near zero. This would mean that the power law model is not suitable to describe these data. Usually, p-values larger than 0.1 are needed to instill confidence in the power law model[30,41]. However, there are many p-values exceeding the significance level of 10%. We can see that p-values increase with increasing threshold, and are mostly concentrated in the range $0.1 \leq q \leq 0.4$. At $q = 1.8$ there are 3 p-values exceeding 50% and at $q = 2$ there are 5 p-values (See Appendix B). The residues pair (49-145) has the largest p-value of 0.815. This pair lies in group G2.

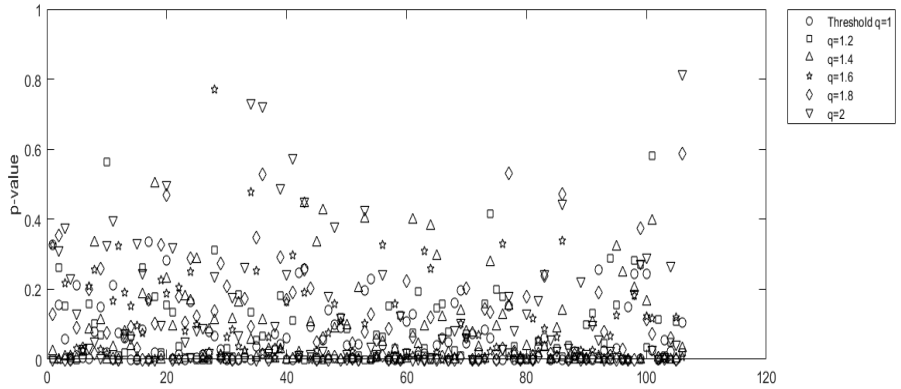



**Fig 3. 3:** The distribution of p-values for 106 pairs of residues at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively.

Table.1 shows the pairs that have p-value greater than 0.5. We can notice that most of the largest p-value lie in group G2 which is shown in bold face. Larger p-values mean the power law model describes the data with high confidence[30].

| pairs | threshold | p-value | α |
|---|---|---|---|
| **118-62** | **2.0** | **0.5** | **1.961** |
| **118-60** | **1.4** | **0.501** | **1.812** |
| **120-62** | **1.8** | **0.528** | **2.078** |
| **125-63** | **1.8** | **0.532** | **1.918** |
| **117-60** | **1.2** | **0.565** | **1.788** |
| **121-59** | **2.0** | **0.575** | **1.922** |
| 49-140 | 1.2 | 0.58 | 1.822 |
| 49-145 | 1.8 | 0.588 | 2.139 |

| | | | |
|---|---|---|---|
| **120-62** | **2.0** | **0.724** | **2.1** |
| **120-60** | **2.0** | **0.733** | **2.282** |
| **119-62** | **1.6** | **0.772** | **2.42** |
| 49-145 | 2.0 | 0.815 | 2.196 |

**Table3. 1**: The pairs of residues that have p-value greater than 0.5.

Fig 3.4 shows the scaling parameter α versus the corresponding p-values at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively. A significant number of p-values are larger than 0.1**.** Thus, many of power models have significant confidence levels. One can also notice that the values of the scaling parameter increase slightly with the p-value. Interestingly, the region with the highest scaling parameter values, are those with the highest p-value.

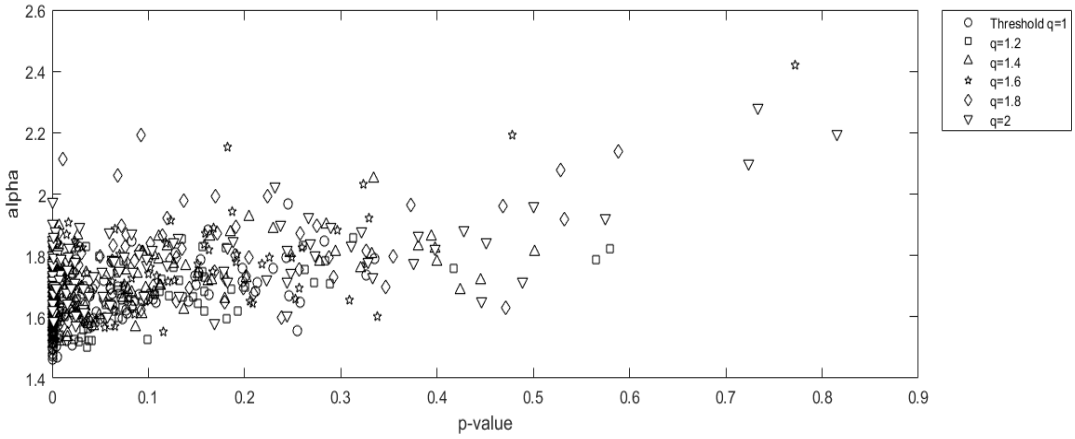



**Fig 3. 4:** The change of p-values with changing the value of α at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively.

Fig 3.5 shows the distribution of the mean recurrence interval $\tau_{avg}$ for 106 pairs of residues at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively.

In general, as the threshold value increases, the mean recurrence interval $\tau_{avg}$of the pairs of residues increases with a systematic shift up. This is expected, because events with large thresholds are extremely rare. The two residue pairs (119-63 & 118-63) have the largest $\tau_{avg}$ at all thresholds and these pairs fall in group G2.

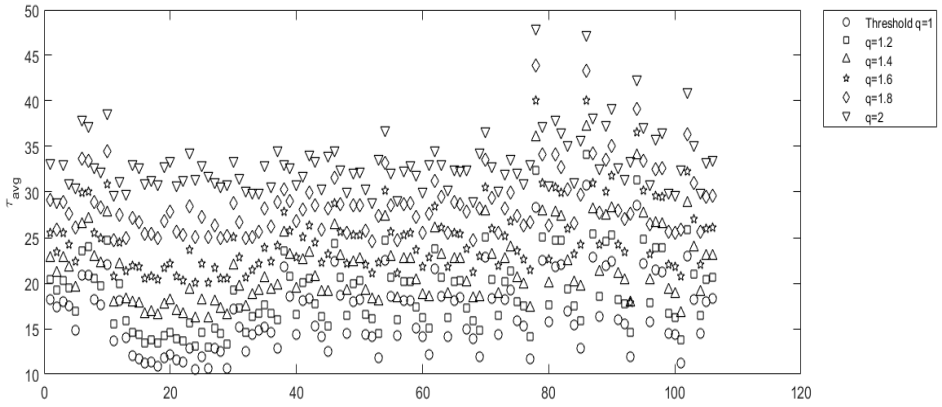



**Fig 3. 5:** The distribution of mean recurrence interval $\boldsymbol{\tau_{avg}}$ for 106 pairs of residues at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively

Fig 3.6 shows the distribution of α-values versus $\boldsymbol{\tau_{avg}}$ , at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively. There is a slight linear increase of the scaling parameter with increasing average recurrence time. The groups belonging to the various thresholds are bunched up in groups that drift towards larger average recurrence times with increasing thresholds.

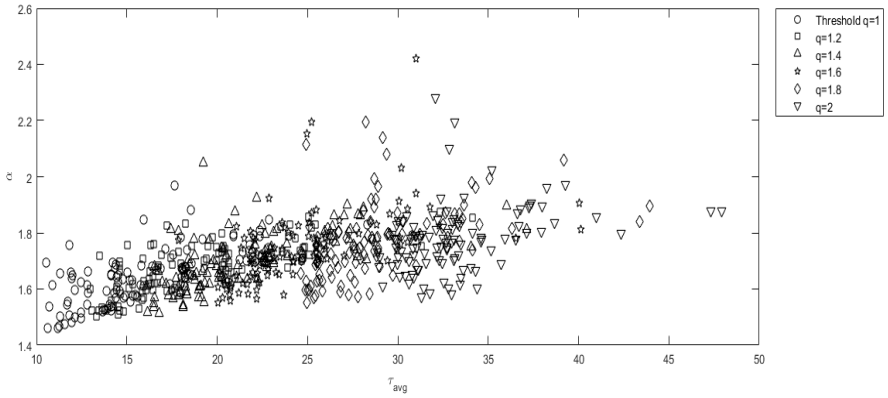



**Fig 3. 6:** The distribution of α-values versus $\boldsymbol{\tau_{avg}}$ at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively.

Fig 3.7 shows the distribution of p-values versus $\tau_{avg}$ at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively. The values are bunched up as groups. This is due to the systematic increase in the mean time for recurrence of very extreme ( highest thresholds).

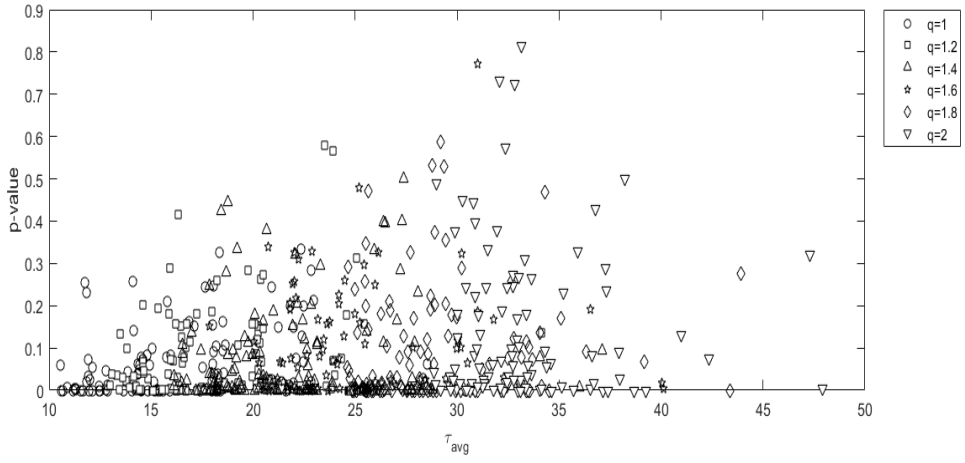



**Fig 3. 7:** The distribution of p-values versus $\boldsymbol{\tau_{avg}}$ at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively.

Fig 3.8 shows the ROC curves for the residue pair 49-145 at thresholds of 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively. The six curves are above the dashed line $D = A(random\ case\ results)$. This is encouraging because it indicates that our predictions are not random. The six curves do not overlap, indicating that the accuracy of this prediction algorithm varies for different pairs of residues. The value of AUC for these residues are 0.0474, 0.0539, 0.05, 0.0597, 0.055, and 0.0638 at thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2.0 respectively.

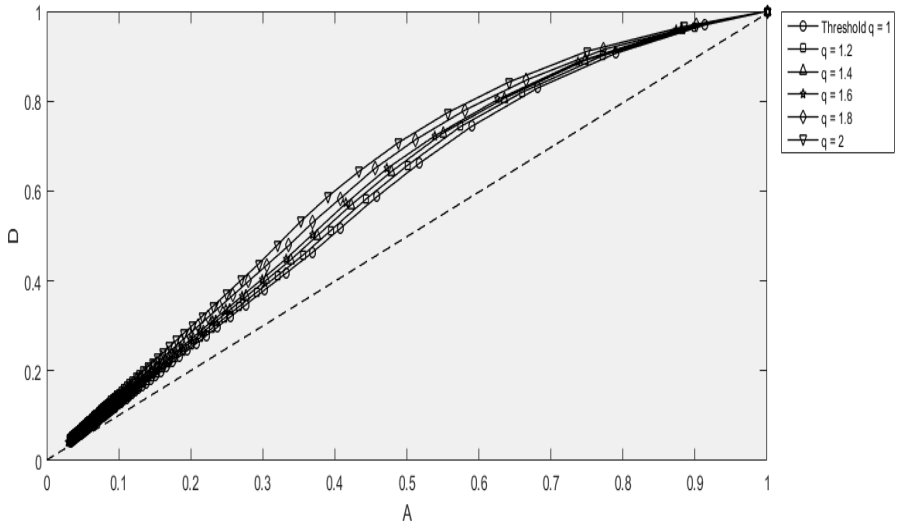



**Fig 3.8:** Plots of ROC curves for residue pair 49-145.

Fig 3.9 shows the distribution of AUC values for all pairs of residues at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2.0 respectively. (Also see Appendix C). In general, the AUC values increase with increasing threshold. Most of the values are above the random prediction value of 0.045; The top three values are 0.0731, 0.0712, 0.0708 for pairs of residues 119-63, 118-63 at

threshold 2.0, and 119-63 at threshold 1.4 respectively. There are 18 values under the random prediction area value of 0.045.

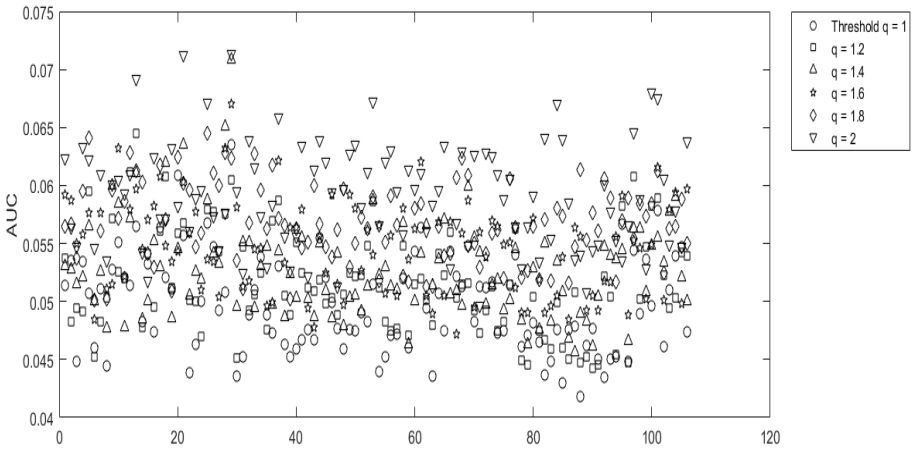



**Fig 3.9** The distribution of AUC values for all pairs of residues at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively.

Fig 3.10 shows the distribution of AUC values versus scaling parameter α at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively. The values are clustered slightly above 0.055. The prediction capability is limited, and the maximum value is well below the optimum value of 0.3.

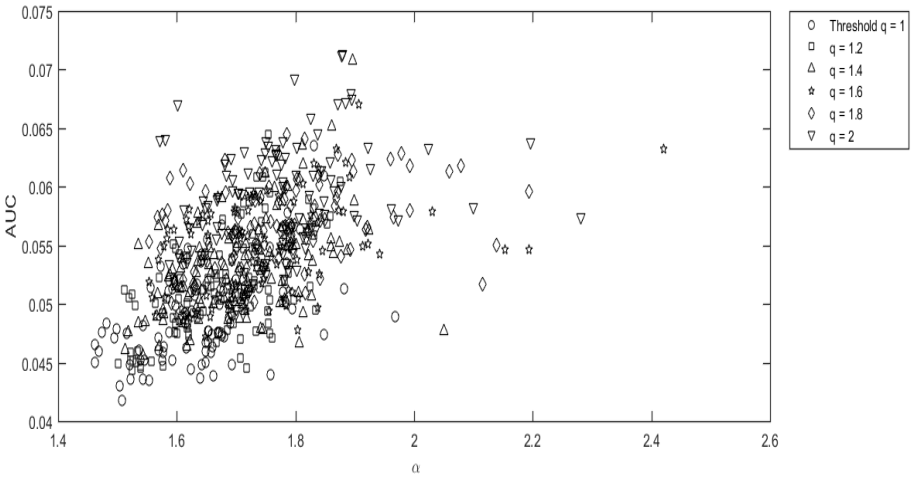



**Fig 3. 10:** The distribution of AUC values versus α at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively.

Fig 3.11 shows the distribution of AUC values versus p-value at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively. A large percentage of the values are below the p value of 0.1. There is a trend of higher AUC values for larger thresholds.

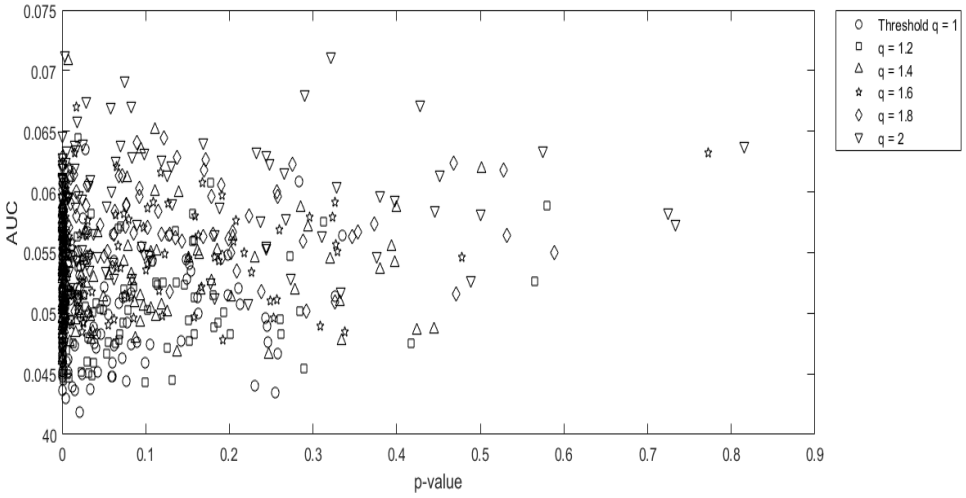



**Fig 3. 11:** The distribution of AUC values versus p-value at positive thresholds 1.0, 1.2, 1.4, 1.6, 1.8, and 2 respectively.

## 4. **Conclusion**

We investigated the dynamics of distance return recurrence intervals in 106 pairs of residues in BLIP protein at different thresholds $q$: $1.0, 1.2, 1.4, 1.6, 1.8, and\ 2.0,$ using recurrence interval analysis. We found that the tails of the recurrence interval distributions obey a power-law model with different significant levels which are independent of the threshold value. Most of the p-values for these power law distributions were larger than the 0.1 significance level. We also looked at the relationship between the power law parameters and the p-values.

The power law pdfs were used to predict extreme events in the distance returns for the residue pairs. This was done using the AUC parameter, which is well known in the ROC curve methodology. Our results showed that the power law models provide predictions which are higher than the random guessing level. However, there is room for improvement. One should in the future investigate other fitting models to the pdfs, in order to iimprove the prediction power of the AUC parameter.

## 5. References

[1] Skliros, A., Zimmermann, M. T., Chakraborty, D., Saraswathi, S., Katebi, A. R., Leelananda, S. P. and Jernigan, R. L. (2012). The importance of slow motions for protein functional loops. *Physical biology*, *9*(1), 014001.

[2] Oiling, L. etc. (2018). Loop. *Foldit Wiki*. foldit.fandom.com/wiki/Loop?oldid=27530.

[3] Shehu,A. and Kavraki,L.E. (2012). Modeling Structures and Motions of Loops in Protein Molecules. *Entropy , 14*, 252-290.

[4] Martin,A.C.R., Toda,K., Stirk,H.J. and Thornton,J.M. (1995). Long loops in proteins. *Protein Eng., 8(11)*, 1093–1101

[5] Yao, P., Dhanik, A., Marz, N., Propper, R., Kou, C., Liu, G., van den Bedem, H., Latombe, J.C., Halperin-Landsberg, I. and Altman, R.B.(2008). Efficient algorithms to explore conformation spaces of flexible protein loops. *IEEE/ACM Trans Comput Biol Bioinform*. *5(4)*:534-545.

[6] Dhar, J. and Chakrabarti,P.(2015). Defining the loop structures in proteins based on composite β-turn mimics. *Protein Engineering, Design & Selection*, *28(6)*, pp. 153–161

[7] Wikipedia contributors. (2016). Beta-lactamase inhibitor protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:22, March 8, 2020. https://en.wikipedia.org/w/index.php?title=Beta lactamase_inhibitor_protein&oldid=728292174

[8] Doran, J. L., Leskiw, B. K., Aippersbach,S. And Jensen, S. E. (1990). Isolation and characterization of a beta-lactamase-inhibitory protein from Streptomyces clavuligerus and cloning and analysis of the corresponding gene. *J. Bacteriol. 172*:4909–4918.

[9] Lim, D., Park, H., De Castro, L. *et al.*(2001). Crystal structure and kinetic analysis of β-lactamase inhibitor protein-II in complex with TEM-1 β-lactamase. *Nat Struct Mol Biol ,8*, 848–852

[10] Strynadka, N., Jensen, S., Alzari, P. *et al.* (1996). A potent new mode of β-lactamase inhibition revealed by the 1.7 Å X-ray crystallographic structure of the TEM-1–BLIP complex. *Nat Struct Mol Biol ,3*, 290–297.

[11] Karain,W.I.(2019). Investigating large-amplitude protein loop motions as extreme events using recurrence interval analysis. *Physica A*,520,1-10.

[12] Chicheportiche, R. and Chakraborti, A.( 2017). A model-free characterization of recurrences in stationary time series. *Physica A*,*474*,312-318.

[13] Chicheportiche, R. and Chakraborti, A.( 2014). Copulas and time series with long-ranged dependencies. *Phys. Rev*. E 89, 042117

[14] Ren,F. and Zhou,W.-X. (2010) . Recurrence interval analysis of trading volumes. *Phys. Rev.* E 81, 066107.

[15] Bunde ,A., Eichner ,J. F., Havlin, S. and Kantelhardt, J. W. (2004) . Return intervals of rare events in records with long-term persistence. *Phys. A*, 342, 308–314

[16] Bunde ,A., Eichner, J. F., Kantelhardt, J. W. and Havlin, S. (2005*).* Long-Term Memory: A Natural Mechanism for the Clustering of Extreme Events and Anomalous Residual Times in Climate Records.*Phys. Rev. Lett.* 94, 048701

[17] Bogachev, M.I., Kireenkov, I. S.,Nifontov, E. M. and Bunde, A.(2009). Statistics of return intervals between long heartbeat intervals and their usability for online prediction of disorders. *New J. Phys.* 11, 063036.

[18] Yamasaki, K., Muchnik, L., Havlin, S., Bunde, A. and Stanley, H. E.(2005). Scaling and memory in volatility return intervals in financial markets, *Proc. Natl. Acad. Sci. U.S.A.* 102, 9424–9428.

[19] Xie, W.-J, Jiang, Z.-Q. and Zhou,W.-X. (2014). Extreme value statistics and recurrence intervals of NYMEX energy futures volatilit. *Econ. Model*. 36 ,8–17.

[20] Bogachev, M. I. and Bunde, A. (2012). Universality in the precipitation and river runoff. *EPL (Europhys. Lett.)* 97 ,48011.

[21] Gallavotti, G .(2008). Fluctuations. *Scholarpedia, 3(6):*5893.

[22] Ren, F. and Zhou, W.-X. (2010) . Recurrence interval analysis of high-frequency financial returns and its application to risk estimation. *New Journal of Physics*, 12, 075030 (16pp)

[23]Epa, V., Winkler, D. and Tran, L. (2012). Adverse Effect of Engineered Nanomaterials.85-96

[24] Pitman, M. R. and Menz, R. I. (2006). Applied Mycology and Biotechnology. 6, Pages 37-59

[25] Berhanu, W. M. and Masunov, A. E.(2014). Polyphenols in Human Health and Disease. 1, pp59-70.

[26] James, C. P., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L. and Schulten K. (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem*, *26(16),* 1781-1802.

[27]Humphrey, W., Dalk, A. and Schulten K. (1996). VMD: Visual Molecular Dynamics. J. *Mol. Graph*. *14(1),* 33-38.

[28] Gretes, M., Lim, D. C., Castro, L. de., Jensen, S. E., Kang, S. G., Lee, K. J. and Strynadka N. C. (2009). Insights into positive and negative requirements for protein-protein interactions by crystallographic analysis of the $\beta$ −lactamase inhibitor protein BLIP, BLIP-I, and BLP. *J. Mol. Biol*. *389(2),* 289-305

[29] Ren, F. and Zhou, W.-X.(2010). Recurrence interval analysis of high-frequency financial returns and its application to risk estimation. *New J. Phys*. *(12) 075030* . doi:10.1088/1367-2630/12/7/075030.

[30] Clauset, A., Shalizi. C. R. and Newman, M. E. J.(2009). Power-law distributions in empirical data. *SIAM Review*, 51, 661–703.

[31] Wichmann, F. A. and Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling and goodness of fit. *Perception & Psychophysics*, 63, 1293–1313.

[32] Muniruzzaman. A. N. M. (1957). On Measures of Location and Dispersion and Tests of Hypotheses in a Pare to Population. *Bulletin of the Calcutta Statistical Association* 7,115.

[33] Clauset, A., Young, M. and Gleditsch. K. S. (2007). On the Frequency of Severe Terrorist Events.*Journal of Conflict Resolution*,51, 58.

[34] Barndorff-Nielsen,O. E. and Cox, D.R. (1995). *Inference and Asymptotics.*

[35] Sornette, D. and Knopoff, L. (1997). The paradox of the expected time until the next earthquake. *Bull. Seim. Soc. Am*, 87, 789-798.

[36] Bogachev, M.I., Eichner, J.F. and Bunde, A. (2007). Effect of nonlinear correlations on the statistics of return intervals in multifractal data sets. *Phys. Rev. Lett*, 99, 240601.

[37] Bogachev, M.I. and Bunde, A. (2011). On the predictability of extreme events in records with linear and nonlinear long-range memory: Efficiency and noise robustness. *Physica A*, 390, 2240-2250.

[38] Jiang,Z.-Q., Canabarro, A.,Podobnik,B., Stanley, H. E. and Zhou, W.-X. (2016) .Early warning of large volatilities based on recurrence interval analysis in Chinese stock markets. *Quantitative Finance*, 16:11, 1713-1724.

[39] Meyer-Baese, A. and Schmid, V. (2014). Statistical and Syntactic Pattern Recognition. *ScienceDirect*. Pages 151-196

https://www.sciencedirect.com/science/article/pii/B9780124095458000066

[40] Lusted, L. B. (1971). Signal Detectability and Medical Decision-Making. *Science*. 171, 3977, pp. 1217-1219.

[41] West, R. (2012). Power Law Behavior of Atmospheric Variability. ll Master Electronic Thesis, Florida State University Libraries

# Appendix A

| Pairs of residues | $\alpha - values$ | | | | | |
|---|---|---|---|---|---|---|
| | Threshold 1.0 | Threshold 1.2 | Threshold 1.4 | Threshold 1.6 | Threshold 1.8 | Threshold 2.0 |
| 116-59 | 1.737948 | 1.693417 | 1.749668 | 1.770944 | 1.793272 | 1.761238 |
| 116-60 | 1.785123 | 1.754398 | 1.740873 | 1.741369 | 1.798809 | 1.83306 |
| 116-61 | 1.641545 | 1.644069 | 1.597675 | 1.773819 | 1.651606 | 1.776129 |
| 116-62 | 1.610041 | 1.639409 | 1.759739 | 1.719565 | 1.68293 | 2.024765 |
| 116-63 | 1.712326 | 1.756095 | 1.779398 | 1.825746 | 1.815218 | 1.858815 |
| 116-64 | 1.468404 | 1.525412 | 1.661941 | 1.61058 | 1.591378 | 1.624458 |
| 116-65 | 1.585336 | 1.61869 | 1.632217 | 1.643674 | 1.695517 | 1.683254 |
| 116-66 | 1.62412 | 1.67505 | 2.049882 | 1.692172 | 1.783906 | 1.748477 |
| 117-59 | 1.705433 | 1.677555 | 1.786487 | 1.700381 | 1.871847 | 1.779052 |
| 117-60 | 1.733733 | 1.787763 | 1.768594 | 1.86924 | 1.830318 | 1.779513 |
| 117-61 | 1.682221 | 1.730805 | 1.745072 | 1.745935 | 1.771404 | 1.822871 |
| 117-62 | 1.69739 | 1.748953 | 1.749679 | 2.031234 | 1.77109 | 1.801052 |
| 117-63 | 1.777942 | 1.75407 | 1.749264 | 1.78013 | 2.059725 | 1.798456 |
| 117-64 | 1.481484 | 1.685007 | 1.576471 | 1.772368 | 1.622809 | 1.640087 |
| 117-65 | 1.572888 | 1.612345 | 1.640066 | 1.651911 | 1.697095 | 1.728659 |
| 117-66 | 1.847822 | 1.833371 | 1.873963 | 1.698391 | 1.70875 | 1.742364 |
| 118-59 | 1.78925 | 1.819823 | 1.704565 | 1.890885 | 1.993252 | 1.973794 |
| 118-60 | 1.712974 | 1.735154 | 1.812312 | 1.785465 | 1.841049 | 1.904695 |
| 118-61 | 1.678874 | 1.754172 | 1.63628 | 1.791907 | 1.81645 | 1.767636 |
| 118-62 | 1.847395 | 1.830838 | 1.887791 | 1.941332 | 1.960066 | 1.961047 |
| 118-63 | 1.791262 | 1.852844 | 1.811431 | 1.812197 | 1.839722 | 1.877518 |
| 118-64 | 1.661943 | 1.563137 | 1.619099 | 1.65194 | 1.648128 | 1.690672 |
| 118-65 | 1.578595 | 1.686957 | 1.657329 | 1.658141 | 1.700055 | 1.668001 |
| 118-66 | 1.672234 | 1.707245 | 1.718765 | 1.794179 | 1.793529 | 1.738535 |
| 119-59 | 1.790274 | 1.702594 | 1.899357 | 1.738798 | 1.786049 | 1.870922 |
| 119-60 | 1.73682 | 1.707919 | 1.714952 | 1.768151 | 1.739954 | 1.805858 |
| 119-61 | 1.613616 | 1.603805 | 1.679377 | 1.692181 | 1.784097 | 1.779386 |
| 119-62 | 1.69831 | 1.858283 | 1.861599 | 2.420441 | 1.978567 | 1.899093 |
| 119-63 | 1.830728 | 1.875266 | 1.895863 | 1.90729 | 1.895112 | 1.879749 |
| 119-64 | 1.542901 | 1.556889 | 1.576464 | 1.62148 | 1.676019 | 1.704012 |
| 119-65 | 1.575771 | 1.598143 | 1.702586 | 1.725744 | 1.641969 | 1.750424 |
| 119-66 | 1.610179 | 1.691192 | 1.71681 | 1.743932 | 1.730387 | 1.750156 |
| 120-59 | 1.778407 | 1.671111 | 1.715618 | 1.845111 | 1.870944 | 1.926671 |
| 120-60 | 1.641561 | 1.664695 | 1.736464 | 2.19349 | 2.193607 | 2.281575 |
| 120-61 | 1.687096 | 1.640131 | 1.654368 | 1.659058 | 1.698565 | 1.736178 |
| 120-62 | 1.675642 | 1.742948 | 1.701458 | 1.784773 | 2.078297 | 2.099604 |
| 120-63 | 1.745396 | 1.721463 | 1.863846 | 1.884721 | 1.848749 | 1.825234 |
| 120-64 | 1.659641 | 1.519765 | 1.705715 | 1.749673 | 1.768118 | 1.667097 |
| 120-65 | 1.592563 | 1.594247 | 1.588733 | 1.585213 | 1.731007 | 1.714314 |
| 120-66 | 1.576127 | 1.594114 | 1.637429 | 1.81916 | 1.746578 | 1.81811 |
| 121-59 | 1.652304 | 1.682705 | 1.794968 | 1.880808 | 1.892287 | 1.922452 |
| 121-60 | 1.669434 | 1.643995 | 1.6976 | 1.755031 | 1.747575 | 1.780175 |
| 121-61 | 1.647956 | 1.662716 | 1.717204 | 1.803541 | 1.75588 | 1.844078 |
| 121-62 | 1.65158 | 1.677743 | 1.800493 | 1.706928 | 1.7931 | 1.778961 |
| 121-63 | 1.699397 | 1.724229 | 1.805209 | 1.839813 | 1.795957 | 1.776797 |
| 121-64 | 1.597929 | 1.757112 | 1.686405 | 1.669042 | 1.692268 | 1.709827 |
| 121-65 | 1.652341 | 1.634294 | 1.644889 | 1.706485 | 1.698911 | 1.715553 |
| 121-66 | 1.657409 | 1.71066 | 1.741022 | 1.83759 | 1.800498 | 1.86484 |
| 122-59 | 1.672282 | 1.66734 | 1.708832 | 1.73159 | 1.717759 | 1.767266 |
| 122-60 | 1.669491 | 1.652675 | 1.695419 | 1.719797 | 1.736596 | 1.754658 |
| 122-61 | 1.69238 | 1.570299 | 1.578164 | 1.822138 | 1.702857 | 1.612039 |
| 122-62 | 1.690337 | 1.726922 | 1.724214 | 1.709775 | 1.774646 | 1.719859 |
| 122-63 | 1.71515 | 1.712673 | 1.777958 | 1.759871 | 1.868561 | 1.884119 |
| 122-64 | 1.758106 | 1.511223 | 1.585466 | 1.61972 | 1.65013 | 1.668163 |
| 122-65 | 1.538279 | 1.597898 | 1.606201 | 1.650047 | 1.65415 | 1.680946 |
| 122-66 | 1.569383 | 1.619577 | 1.605514 | 1.781189 | 1.741086 | 1.713505 |
| 123-59 | 1.680135 | 1.732819 | 1.71614 | 1.781478 | 1.782628 | 1.70327 |
| 123-60 | 1.586673 | 1.699637 | 1.63428 | 1.872885 | 1.841068 | 1.812947 |
| 123-61 | 1.64888 | 1.760588 | 1.570037 | 1.728321 | 1.627179 | 1.84176 |
| 123-62 | 1.64711 | 1.652712 | 1.709981 | 1.706077 | 1.992732 | 1.693994 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 123-63 | 1.807905 | 1.781811 | 1.822314 | 1.831249 | 1.832754 | 1.768977 |
| 123-64 | 1.625354 | 1.617256 | 1.568905 | 1.712951 | 1.673397 | 1.675698 |
| 123-65 | 1.522156 | 1.617792 | 1.607644 | 1.656027 | 1.628267 | 1.648792 |
| 123-66 | 1.630651 | 1.615471 | 1.828821 | 1.82538 | 1.921008 | 1.72333 |
| 124-59 | 1.676014 | 1.666119 | 1.812762 | 1.804498 | 1.771655 | 1.802386 |
| 124-60 | 1.641884 | 1.727791 | 1.827458 | 1.837746 | 1.876856 | 1.787994 |
| 124-61 | 1.594994 | 1.664506 | 1.620894 | 1.64885 | 1.588535 | 1.643338 |
| 124-62 | 1.882144 | 1.657707 | 1.689057 | 1.706533 | 1.681377 | 1.748519 |
| 124-63 | 1.756392 | 1.731185 | 1.764654 | 1.796568 | 1.838685 | 1.816664 |
| 124-64 | 1.614581 | 1.628119 | 1.688132 | 1.716016 | 1.734553 | 1.783086 |
| 124-65 | 1.592373 | 1.619737 | 1.662548 | 1.624146 | 1.664669 | 1.700013 |
| 124-66 | 1.632512 | 1.575984 | 1.633336 | 1.802482 | 1.720567 | 1.767061 |
| 125-59 | 1.634738 | 1.684349 | 1.611691 | 1.637273 | 1.692266 | 1.692134 |
| 125-60 | 1.676353 | 1.756209 | 1.777924 | 1.782206 | 1.821368 | 1.829262 |
| 125-61 | 1.515884 | 1.695914 | 1.813017 | 1.578303 | 1.595941 | 1.64499 |
| 125-62 | 1.556204 | 1.824514 | 1.773662 | 1.922413 | 1.80614 | 1.693677 |
| 125-63 | 1.742143 | 1.699255 | 1.740058 | 1.7874 | 1.918335 | 1.826714 |
| 125-64 | 1.536474 | 1.500922 | 1.546629 | 1.598782 | 1.610631 | 1.746192 |
| 125-65 | 1.495116 | 1.538437 | 1.617625 | 1.615179 | 1.628706 | 1.621483 |
| 125-66 | 1.542726 | 1.602934 | 1.592437 | 1.616152 | 1.723675 | 1.669048 |
| 48-133 | 1.461708 | 1.600462 | 1.517579 | 1.552987 | 1.596023 | 1.586747 |
| 48-134 | 1.63882 | 1.666253 | 1.534812 | 1.564252 | 1.580046 | 1.580916 |
| 48-135 | 1.692249 | 1.522407 | 1.611038 | 1.608006 | 1.596161 | 1.603751 |
| 48-136 | 1.587291 | 1.530162 | 1.551397 | 1.579001 | 1.584106 | 1.601953 |
| 48-137 | 1.503969 | 1.53364 | 1.567104 | 1.569495 | 1.567795 | 1.570971 |
| 48-138 | 1.473306 | 1.530735 | 1.62304 | 1.602612 | 1.630146 | 1.652562 |
| 48-139 | 1.497891 | 1.524116 | 1.542218 | 1.595344 | 1.57251 | 1.616547 |
| 48-140 | 1.507181 | 1.539123 | 1.566153 | 1.583465 | 1.610357 | 1.604616 |
| 48-141 | 1.615511 | 1.521775 | 1.534833 | 1.617075 | 1.552697 | 1.720657 |
| 48-142 | 1.653921 | 1.524761 | 1.512897 | 1.559079 | 1.575454 | 1.796359 |
| 48-143 | 1.461371 | 1.717953 | 1.680293 | 1.63299 | 1.686828 | 1.791398 |
| 48-144 | 1.55358 | 1.605879 | 1.637639 | 1.692199 | 1.801884 | 1.876772 |
| 48-145 | 1.647972 | 1.583548 | 1.608572 | 1.699667 | 1.780955 | 1.84999 |
| 49-133 | 1.54552 | 1.707032 | 1.635792 | 1.746957 | 2.115436 | 1.77955 |
| 49-134 | 1.675325 | 1.748497 | 1.756631 | 1.719946 | 1.757838 | 1.783208 |
| 49-135 | 1.520793 | 1.576608 | 1.806919 | 1.622366 | 1.726906 | 1.794893 |
| 49-136 | 1.732443 | 1.744569 | 1.738058 | 1.793012 | 1.829639 | 1.83796 |
| 49-137 | 1.96818 | 1.783466 | 1.923601 | 2.153434 | 1.792422 | 1.848617 |
| 49-138 | 1.640304 | 1.71293 | 1.714622 | 1.776176 | 1.965113 | 1.802025 |
| 49-139 | 1.794344 | 1.829755 | 1.859738 | 1.913096 | 1.866322 | 1.893082 |
| 49-140 | 1.786663 | 1.822406 | 1.860062 | 1.838263 | 1.898034 | 1.89446 |
| 49-141 | 1.568667 | 1.721987 | 1.654182 | 1.684984 | 1.815196 | 1.859832 |
| 49-142 | 1.720424 | 1.753025 | 1.63517 | 1.696123 | 1.736824 | 1.789792 |
| 49-143 | 1.61692 | 1.797523 | 1.675916 | 1.728744 | 1.783783 | 1.840252 |
| 49-144 | 1.618271 | 1.671928 | 1.684083 | 1.713791 | 1.73156 | 1.732577 |
| 49-145 | 1.658568 | 1.692503 | 1.689289 | 1.752117 | 2.139306 | 2.196251 |

# Appendix B

| Pairs of residues | p-values | | | | | |
|---|---|---|---|---|---|---|
| | Threshold 1.0 | Threshold 1.2 | Threshold 1.4 | Threshold 1.6 | Threshold 1.8 | Threshold 2.0 |
| 116-59 | 0.326 | 1.00E-03 | 0.023 | 0.326 | 0.129 | 0 |
| 116-60 | 0.154 | 0.262 | 0 | 0.00E+00 | 0.354 | 0.311 |
| 116-61 | 0.059 | 0.151 | 7.00E-03 | 0.218 | 0.00E+00 | 0.376 |
| 116-62 | 0.00E+00 | 0.003 | 0.006 | 1.20E-02 | 0.00E+00 | 0.232 |
| 116-63 | 0.213 | 0 | 1.00E-03 | 0.031 | 0.09 | 0.131 |
| 116-64 | 5.00E-03 | 2.30E-02 | 0.023 | 0.024 | 3.20E-02 | 3.20E-02 |
| 116-65 | 1.00E-03 | 0.158 | 8.40E-02 | 2.08E-01 | 0.199 | 3.40E-02 |
| 116-66 | 0.077 | 0.102 | 0.334 | 0.257 | 0.015 | 0.052 |
| 117-59 | 0.149 | 7.00E-02 | 0.111 | 0.027 | 0.258 | 0.003 |
| 117-60 | 0.019 | 0.565 | 0 | 0.015 | 0.076 | 0.328 |
| 117-61 | 0.211 | 0 | 0.035 | 0.167 | 1.00E-03 | 0.398 |
| 117-62 | 0.076 | 0.009 | 0 | 0.323 | 0.00E+00 | 0 |
| 117-63 | 0.061 | 0.019 | 0.078 | 0.191 | 0.068 | 0.075 |
| 117-64 | 0.00E+00 | 0.068 | 3.70E-02 | 0.151 | 9.20E-02 | 6.00E-02 |
| 117-65 | 0.00E+00 | 0.00E+00 | 3.00E-03 | 0.097 | 0.03 | 0.333 |
| 117-66 | 0.083 | 0.261 | 0.185 | 0.074 | 0.00E+00 | 0.248 |
| 118-59 | 0.335 | 0 | 0 | 0.168 | 0.17 | 0 |
| 118-60 | 0.018 | 0.178 | 0.501 | 1.00E-03 | 0.102 | 0 |
| 118-61 | 0.009 | 0.018 | 0 | 0.226 | 0.326 | 0.098 |
| 118-62 | 0.283 | 0.156 | 0.23 | 0.187 | 0.468 | 0.5 |
| 118-63 | 1.00E-03 | 0.135 | 0.094 | 0.003 | 0 | 0.321 |
| 118-64 | 0.012 | 3.70E-02 | 0.00E+00 | 0.205 | 0.179 | 0.002 |
| 118-65 | 0.00E+00 | 0.003 | 0.179 | 0.08 | 0.105 | 0.053 |
| 118-66 | 0.163 | 0.005 | 0.166 | 0.249 | 0.288 | 8.00E-03 |
| 119-59 | 0.004 | 0.00E+00 | 0.285 | 0.1 | 0.122 | 0.083 |
| 119-60 | 0.013 | 0 | 1.00E-03 | 0.005 | 0.01 | 0.09 |
| 119-61 | 0.078 | 0.00E+00 | 0.083 | 1.00E-03 | 0.017 | 0.018 |
| 119-62 | 0.068 | 0.313 | 0.111 | 0.772 | 0.137 | 0.237 |
| 119-63 | 0.028 | 0.003 | 0.007 | 0.017 | 0.275 | 0.003 |
| 119-64 | 0 | 3.20E-02 | 0.00E+00 | 6.40E-02 | 0.209 | 0.007 |
| 119-65 | 0.00E+00 | 0.007 | 0.113 | 0.083 | 0.00E+00 | 0.178 |
| 119-66 | 0.00E+00 | 0.186 | 0.162 | 0.036 | 1.00E-03 | 0.07 |
| 120-59 | 0.003 | 0.004 | 0.016 | 0.024 | 0.172 | 0.266 |
| 120-60 | 0 | 0.136 | 0.056 | 0.478 | 0.092 | 0.733 |
| 120-61 | 0.038 | 5.56E-02 | 0.003 | 0.253 | 3.47E-01 | 0 |
| 120-62 | 0.015 | 0.067 | 0.112 | 0.02 | 0.528 | 0.724 |
| 120-63 | 0.021 | 1.00E-03 | 0.027 | 0.065 | 0 | 0.018 |
| 120-64 | 0.072 | 2.80E-02 | 0.002 | 0.003 | 0.028 | 0.095 |
| 120-65 | 0.043 | 1.81E-01 | 2.90E-02 | 0.029 | 0.292 | 0.489 |
| 120-66 | 0.06 | 0.00E+00 | 0.00E+00 | 0.163 | 0.169 | 0.244 |
| 121-59 | 0 | 0.112 | 0.002 | 0.296 | 0.191 | 0.575 |
| 121-60 | 0.246 | 0.003 | 0.02 | 0.005 | 0.006 | 0.014 |
| 121-61 | 0.258 | 0 | 0.445 | 0.192 | 0.257 | 0.451 |
| 121-62 | 0.094 | 0.008 | 0.106 | 0 | 0.204 | 0.015 |
| 121-63 | 0 | 0.003 | 0.332 | 0.025 | 0.05 | 0 |
| 121-64 | 0.055 | 0 | 0.424 | 1.00E-03 | 0.06 | 0.024 |
| 121-65 | 0.142 | 1.00E-03 | 0.007 | 0.075 | 0 | 0.182 |
| 121-66 | 0.099 | 1.00E-03 | 0.088 | 0.158 | 0.006 | 0.38 |
| 122-59 | 0.09 | 0.121 | 0.00E+00 | 1.09E-01 | 0 | 0.119 |
| 122-60 | 0.039 | 0.086 | 9.70E-02 | 1.00E-03 | 1.00E-03 | 1.60E-02 |
| 122-61 | 1.00E-03 | 0.003 | 1.00E-03 | 0.022 | 0.045 | 1.00E-03 |
| 122-62 | 0.042 | 0.011 | 0.202 | 0.002 | 0 | 0 |
| 122-63 | 0.198 | 0.00E+00 | 0.4 | 0.102 | 0.077 | 0.428 |
| 122-64 | 0.23 | 0.00E+00 | 1.00E-03 | 0.00E+00 | 1.29E-01 | 0.038 |
| 122-65 | 7.00E-03 | 6.90E-02 | 0.004 | 0.004 | 5.40E-02 | 0.015 |
| 122-66 | 0.041 | 0.064 | 4.90E-02 | 0.327 | 0.01 | 0.244 |
| 123-59 | 0 | 0.152 | 0.003 | 1.00E-03 | 0.086 | 0.00E+00 |
| 123-60 | 0 | 0.032 | 2.80E-02 | 0.159 | 0.005 | 0.006 |
| 123-61 | 0 | 0.126 | 0.00E+00 | 0 | 0.00E+00 | 0.126 |
| 123-62 | 1.08E-01 | 0.02 | 1.00E-03 | 0.008 | 0.224 | 0.02 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 123-63 | 0.128 | 0.074 | 0.398 | 0.009 | 0.063 | 0 |
| 123-64 | 0.01 | 0.193 | 4.30E-02 | 0.003 | 0.02 | 1.00E-03 |
| 123-65 | 1.50E-02 | 0.031 | 9.30E-02 | 0.309 | 0.00E+00 | 1.00E-03 |
| 123-66 | 1.00E-03 | 0.00E+00 | 0.38 | 0.26 | 0.119 | 7.00E-03 |
| 124-59 | 0.039 | 0.146 | 0.294 | 0.018 | 0 | 0.005 |
| 124-60 | 0.062 | 0.158 | 0.119 | 0 | 0.007 | 0.062 |
| 124-61 | 0.05 | 0.079 | 0.019 | 0 | 0.00E+00 | 0 |
| 124-62 | 0.162 | 0.026 | 0 | 0.006 | 0.00E+00 | 0.084 |
| 124-63 | 0.198 | 0.01 | 0.139 | 0.102 | 0.004 | 0.104 |
| 124-64 | 0.046 | 0.073 | 0.079 | 0.062 | 0.202 | 0.064 |
| 124-65 | 0.073 | 0 | 0.074 | 0.00E+00 | 0.00E+00 | 0.002 |
| 124-66 | 0.019 | 0.00E+00 | 3.40E-02 | 0.07 | 0 | 0 |
| 125-59 | 0.083 | 0.158 | 2.00E-03 | 1.00E-03 | 0.143 | 0.004 |
| 125-60 | 0.062 | 0.417 | 0.278 | 0.021 | 0.135 | |
| 125-61 | 0.00E+00 | 0.201 | 0.093 | 3.40E-02 | 0 | 0.016 |
| 125-62 | 0.00E+00 | 0.18 | 0.126 | 0.329 | 0.03 | 0.032 |
| 125-63 | 0.152 | 0.018 | 0.151 | 0.006 | 0.532 | 0.182 |
| 125-64 | 0 | 3.60E-02 | 1.60E-02 | 2.00E-03 | 2.10E-02 | 0.084 |
| 125-65 | 0.00E+00 | 1.00E-03 | 0 | 0.011 | 0.002 | 0 |
| 125-66 | 1.00E-03 | 0 | 0.038 | 0 | 0.18 | 0.132 |
| 48-133 | 0.00E+00 | 0.013 | 9.00E-03 | 1.16E-01 | 0 | 1.00E-03 |
| 48-134 | 0.034 | 0.053 | 1.60E-02 | 5.50E-02 | 8.00E-03 | 0.169 |
| 48-135 | 0.06 | 3.80E-02 | 0.034 | 0.086 | 0.238 | 0.244 |
| 48-136 | 0 | 1.00E-03 | 0.00E+00 | 0.00E+00 | 1.70E-02 | 0.058 |
| 48-137 | 4.00E-03 | 0.03 | 8.70E-02 | 0.065 | 0.006 | 0.025 |
| 48-138 | 0.00E+00 | 6.00E-03 | 0.137 | 0.338 | 0.471 | 0.446 |
| 48-139 | 0.00E+00 | 4.10E-02 | 1.00E-03 | 0.00E+00 | 0.00E+00 | 0.033 |
| 48-140 | 2.10E-02 | 8.00E-03 | 0.00E+00 | 6.00E-03 | 0.014 | 2.50E-02 |
| 48-141 | 0.007 | 2.00E-03 | 0.00E+00 | 0.026 | 0.00E+00 | 0.223 |
| 48-142 | 0.002 | 9.90E-02 | 0.00E+00 | 1.00E-03 | 1.00E-03 | 0.017 |
| 48-143 | 0.00E+00 | 0.132 | 0.106 | 0.025 | 0.004 | 0.098 |
| 48-144 | 0.255 | 1.00E-03 | 1.00E-03 | 0.013 | 0.19 | 0.059 |
| 48-145 | 0.024 | 4.50E-02 | 0 | 1.00E-03 | 0.077 | 0 |
| 49-133 | 0.00E+00 | 0.289 | 0 | 0.067 | 0.011 | 0.024 |
| 49-134 | 0.031 | 0.156 | 0.32 | 0.127 | 0 | 0.008 |
| 49-135 | 0.00E+00 | 1.00E-03 | 0.247 | 0 | 0.029 | 1.00E-03 |
| 49-136 | 0.148 | 0 | 0 | 0 | 0 | 0 |
| 49-137 | 0.245 | 0.284 | 0.204 | 0.182 | 0 | 0.188 |
| 49-138 | 0.002 | 0.272 | 0 | 0 | 0.373 | 0.273 |
| 49-139 | 0.243 | 0.035 | 0.164 | 0.123 | 0.115 | 0.29 |
| 49-140 | 0.025 | 0.58 | 0.394 | 0.118 | 0.072 | 0.029 |
| 49-141 | 0 | 0.114 | 0.005 | 0 | 0.003 | 1.00E-03 |
| 49-142 | 0.0536 | 0.0108 | 4.00E-04 | 0.0044 | 1.20E-03 | 0.0088 |
| 49-143 | 2.20E-02 | 0.049 | 0 | 0.004 | 0.062 | 0.268 |
| 49-144 | 0.00E+00 | 0.107 | 1.00E-03 | 0.12 | 0 | 0.043 |
| 49-145 | 0.105 | 0.019 | 3.10E-02 | 0.006 | 0.588 | 0.815 |

# Appendix C

| Pairs of residues | AUC | | | | | |
|---|---|---|---|---|---|---|
| | Threshold 1.0 | Threshold 1.2 | Threshold 1.4 | Threshold 1.6 | Threshold 1.8 | Threshold 2.0 |
| 116-59 | 0.0514 | 0.0537 | 0.0531 | 0.0592 | 0.0565 | 0.0623 |
| 116-60 | 0.0535 | 0.0483 | 0.0528 | 0.0587 | 0.0567 | 0.0564 |
| 116-61 | 0.0448 | 0.0494 | 0.0515 | 0.055 | 0.0537 | 0.0547 |
| 116-62 | 0.0534 | 0.0491 | 0.0521 | 0.0558 | 0.0595 | 0.0633 |
| 116-63 | 0.0507 | 0.0595 | 0.0565 | 0.0576 | 0.0641 | 0.0622 |
| 116-64 | 0.046 | 0.0452 | 0.05 | 0.0485 | 0.0501 | 0.0546 |
| 116-65 | 0.0511 | 0.0483 | 0.0526 | 0.0576 | 0.0561 | 0.061 |
| 116-66 | 0.0444 | 0.0503 | 0.0477 | 0.0511 | 0.0502 | 0.0535 |
| 117-59 | 0.0528 | 0.0572 | 0.06 | 0.0515 | 0.0596 | 0.0601 |
| 117-60 | 0.0551 | 0.0526 | 0.0584 | 0.0632 | 0.0572 | 0.0605 |
| 117-61 | 0.0521 | 0.052 | 0.0478 | 0.0522 | 0.0586 | 0.0593 |
| 117-62 | 0.0514 | 0.0612 | 0.0572 | 0.0579 | 0.0628 | 0.061 |
| 117-63 | 0.0565 | 0.0645 | 0.0611 | 0.0597 | 0.0613 | 0.0692 |
| 117-64 | 0.0483 | 0.0478 | 0.0485 | 0.0545 | 0.0603 | 0.0542 |
| 117-65 | 0.0532 | 0.0541 | 0.05 | 0.0571 | 0.0541 | 0.0518 |
| 117-66 | 0.0474 | 0.0495 | 0.0552 | 0.0582 | 0.053 | 0.0624 |
| 118-59 | 0.0564 | 0.0561 | 0.053 | 0.0608 | 0.0618 | 0.0572 |
| 118-60 | 0.0521 | 0.0608 | 0.0619 | 0.0548 | 0.0571 | 0.0572 |
| 118-61 | 0.0511 | 0.0512 | 0.0486 | 0.0534 | 0.0509 | 0.0632 |
| 118-62 | 0.0609 | 0.0559 | 0.0545 | 0.0543 | 0.0624 | 0.0582 |
| 118-63 | 0.0557 | 0.0568 | 0.0635 | 0.0603 | 0.0602 | 0.0712 |
| 118-64 | 0.0439 | 0.0501 | 0.0505 | 0.0559 | 0.0596 | 0.0561 |
| 118-65 | 0.0463 | 0.05 | 0.0526 | 0.0577 | 0.0547 | 0.0589 |
| 118-66 | 0.05 | 0.047 | 0.0519 | 0.051 | 0.0559 | 0.0596 |
| 119-59 | 0.0568 | 0.0579 | 0.0587 | 0.0536 | 0.0645 | 0.0671 |
| 119-60 | 0.0547 | 0.0577 | 0.0541 | 0.0534 | 0.0611 | 0.0574 |
| 119-61 | 0.0492 | 0.0543 | 0.0532 | 0.0504 | 0.06 | 0.0544 |
| 119-62 | 0.0508 | 0.0575 | 0.0651 | 0.0632 | 0.0628 | 0.0576 |
| 119-63 | 0.0635 | 0.0605 | 0.0708 | 0.067 | 0.0623 | 0.0713 |
| 119-64 | 0.0436 | 0.0451 | 0.0495 | 0.0581 | 0.0535 | 0.0595 |
| 119-65 | 0.0452 | 0.0547 | 0.0551 | 0.0513 | 0.0511 | 0.0525 |
| 119-66 | 0.0488 | 0.0492 | 0.0551 | 0.0518 | 0.0568 | 0.0639 |
| 120-59 | 0.051 | 0.0506 | 0.0494 | 0.0545 | 0.0627 | 0.0616 |
| 120-60 | 0.0538 | 0.0525 | 0.0531 | 0.0546 | 0.0596 | 0.0574 |
| 120-61 | 0.0488 | 0.0476 | 0.0547 | 0.0496 | 0.0563 | 0.053 |
| 120-62 | 0.0473 | 0.057 | 0.0497 | 0.0499 | 0.0618 | 0.0583 |
| 120-63 | 0.0531 | 0.0587 | 0.0543 | 0.0621 | 0.0572 | 0.0659 |
| 120-64 | 0.0463 | 0.0505 | 0.0487 | 0.0533 | 0.0566 | 0.0556 |
| 120-65 | 0.0452 | 0.0488 | 0.0524 | 0.0564 | 0.0502 | 0.0527 |
| 120-66 | 0.0459 | 0.0551 | 0.0503 | 0.0564 | 0.0563 | 0.0556 |
| 121-59 | 0.0467 | 0.0525 | 0.0555 | 0.0579 | 0.0546 | 0.0634 |
| 121-60 | 0.0476 | 0.0512 | 0.0502 | 0.0494 | 0.0551 | 0.0522 |
| 121-61 | 0.0467 | 0.0509 | 0.0487 | 0.0478 | 0.06 | 0.0614 |
| 121-62 | 0.0555 | 0.0519 | 0.055 | 0.0556 | 0.0567 | 0.0639 |
| 121-63 | 0.0521 | 0.0498 | 0.0509 | 0.0525 | 0.0547 | 0.062 |
| 121-64 | 0.0499 | 0.0503 | 0.0486 | 0.0592 | 0.0538 | 0.0594 |
| 121-65 | 0.0477 | 0.0488 | 0.0541 | 0.0515 | 0.0562 | 0.0513 |
| 121-66 | 0.0459 | 0.0506 | 0.0479 | 0.0497 | 0.0597 | 0.0597 |
| 122-59 | 0.0476 | 0.0525 | 0.0503 | 0.0592 | 0.053 | 0.0627 |
| 122-60 | 0.0475 | 0.0522 | 0.0505 | 0.058 | 0.055 | 0.0635 |
| 122-61 | 0.049 | 0.0523 | 0.0491 | 0.0527 | 0.0573 | 0.0581 |
| 122-62 | 0.0483 | 0.0548 | 0.0513 | 0.0562 | 0.0563 | 0.0612 |
| 122-63 | 0.0515 | 0.0586 | 0.0587 | 0.054 | 0.059 | 0.0672 |
| 122-64 | 0.044 | 0.0512 | 0.0523 | 0.0566 | 0.0518 | 0.0566 |
| 122-65 | 0.0452 | 0.0483 | 0.053 | 0.0507 | 0.0551 | 0.062 |
| 122-66 | 0.0471 | 0.0475 | 0.0513 | 0.0557 | 0.0591 | 0.063 |
| 123-59 | 0.0472 | 0.0477 | 0.051 | 0.0505 | 0.0566 | 0.0595 |
| 123-60 | 0.052 | 0.0548 | 0.0568 | 0.058 | 0.0522 | 0.0558 |
| 123-61 | 0.046 | 0.0471 | 0.0463 | 0.0536 | 0.0517 | 0.0614 |
| 123-62 | 0.0543 | 0.0515 | 0.05 | 0.0563 | 0.058 | 0.0597 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 123-63 | 0.0564 | 0.052 | 0.0541 | 0.062 | 0.0584 | 0.0611 |
| 123-64 | 0.0494 | 0.0501 | 0.0566 | 0.0505 | 0.0513 | 0.0561 |
| 123-65 | 0.0436 | 0.048 | 0.0514 | 0.0489 | 0.0513 | 0.0583 |
| 123-66 | 0.0507 | 0.0541 | 0.0536 | 0.0569 | 0.0565 | 0.0596 |
| 124-59 | 0.0475 | 0.0523 | 0.0571 | 0.0505 | 0.0539 | 0.0634 |
| 124-60 | 0.0543 | 0.056 | 0.0506 | 0.0569 | 0.0541 | 0.0529 |
| 124-61 | 0.0522 | 0.0503 | 0.0509 | 0.0472 | 0.0607 | 0.0568 |
| 124-62 | 0.0513 | 0.0515 | 0.0496 | 0.0559 | 0.0623 | 0.0629 |
| 124-63 | 0.0548 | 0.0516 | 0.0599 | 0.0587 | 0.0609 | 0.0548 |
| 124-64 | 0.0483 | 0.0492 | 0.0555 | 0.0495 | 0.055 | 0.0626 |
| 124-65 | 0.0513 | 0.0473 | 0.0542 | 0.056 | 0.053 | 0.0496 |
| 124-66 | 0.0512 | 0.0492 | 0.0497 | 0.0538 | 0.0541 | 0.0628 |
| 125-59 | 0.0517 | 0.0513 | 0.0514 | 0.057 | 0.0559 | 0.0625 |
| 125-60 | 0.0473 | 0.0475 | 0.0519 | 0.0565 | 0.0562 | 0.061 |
| 125-61 | 0.0475 | 0.0483 | 0.0492 | 0.0549 | 0.052 | 0.0588 |
| 125-62 | 0.0514 | 0.0515 | 0.0501 | 0.0551 | 0.0606 | 0.0606 |
| 125-63 | 0.054 | 0.0563 | 0.0545 | 0.0512 | 0.0564 | 0.0565 |
| 125-64 | 0.0461 | 0.0449 | 0.0484 | 0.049 | 0.0521 | 0.0529 |
| 125-65 | 0.0471 | 0.0445 | 0.0463 | 0.049 | 0.0527 | 0.0565 |
| 125-66 | 0.0482 | 0.0504 | 0.0529 | 0.0572 | 0.0565 | 0.0591 |
| 48-133 | 0.0465 | 0.0475 | 0.0476 | 0.0519 | 0.0519 | 0.0534 |
| 48-134 | 0.0437 | 0.0467 | 0.055 | 0.049 | 0.0571 | 0.0641 |
| 48-135 | 0.0448 | 0.0459 | 0.0483 | 0.0496 | 0.0518 | 0.0554 |
| 48-136 | 0.0476 | 0.0499 | 0.0534 | 0.0505 | 0.058 | 0.067 |
| 48-137 | 0.043 | 0.046 | 0.0508 | 0.0538 | 0.0574 | 0.064 |
| 48-138 | 0.0476 | 0.045 | 0.0468 | 0.0485 | 0.0516 | 0.0585 |
| 48-139 | 0.0479 | 0.0508 | 0.0456 | 0.0564 | 0.0547 | 0.0536 |
| 48-140 | 0.0418 | 0.0447 | 0.0489 | 0.0556 | 0.0614 | 0.054 |
| 48-141 | 0.0462 | 0.0452 | 0.0483 | 0.0492 | 0.0553 | 0.0508 |
| 48-142 | 0.0477 | 0.0443 | 0.0461 | 0.0505 | 0.0576 | 0.0544 |
| 48-143 | 0.045 | 0.0445 | 0.0504 | 0.0492 | 0.0561 | 0.0549 |
| 48-144 | 0.0435 | 0.0522 | 0.0547 | 0.0518 | 0.0606 | 0.0601 |
| 48-145 | 0.045 | 0.0504 | 0.0539 | 0.0517 | 0.0589 | 0.0577 |
| 49-133 | 0.0452 | 0.0454 | 0.0536 | 0.0556 | 0.0517 | 0.0553 |
| 49-134 | 0.0567 | 0.0582 | 0.0544 | 0.0591 | 0.057 | 0.0535 |
| 49-135 | 0.0448 | 0.0447 | 0.0466 | 0.0488 | 0.0589 | 0.0566 |
| 49-136 | 0.0544 | 0.0608 | 0.0563 | 0.0553 | 0.0554 | 0.0646 |
| 49-137 | 0.0489 | 0.0502 | 0.0563 | 0.0546 | 0.0584 | 0.0588 |
| 49-138 | 0.0536 | 0.0547 | 0.0508 | 0.0502 | 0.0574 | 0.0529 |
| 49-139 | 0.0496 | 0.0582 | 0.0548 | 0.0549 | 0.0584 | 0.068 |
| 49-140 | 0.0578 | 0.0589 | 0.0555 | 0.0616 | 0.0613 | 0.0675 |
| 49-141 | 0.0461 | 0.0523 | 0.0533 | 0.0501 | 0.0525 | 0.0606 |
| 49-142 | 0.051 | 0.0494 | 0.0577 | 0.0579 | 0.0563 | 0.0549 |
| 49-143 | 0.0539 | 0.0503 | 0.059 | 0.0594 | 0.0565 | 0.0578 |
| 49-144 | 0.0547 | 0.0542 | 0.0521 | 0.0498 | 0.0588 | 0.0547 |
| 49-145 | 0.0474 | 0.0539 | 0.05 | 0.0597 | 0.055 | 0.0638 |