Influence of the Active Site Flexibility on the Efficiency of Substrate Activation in the Active Sites of Bi-Zinc Metallo-β-Lactamases

Alexandra V. Krivitskaya 1 and Maria G. Khrenova 1,2,*

1 Bach Institute of Biochemistry, Federal Research Centre “Fundamentals of Biotechnology” of the Russian Academy of Sciences, 119071 Moscow, Russia
2 Department of Chemistry, Interdisciplinary Scientific and Educational School of Moscow University «Brain, Cognitive Systems, Artificial Intelligence», Lomonosov Moscow State University, 119991 Moscow, Russia
* Correspondence: mkhrenova@lcc.chem.msu.ru

Abstract: The influence of the active site flexibility on the efficiency of catalytic reaction is studied by taking two members of metallo-β-lactamases, L1 and NDM-1, with the same substrate, imipenem. Active sites of these proteins are covered by L10 loops, and differences in their amino acid compositions affect their rigidity. A more flexible loop in the NDM-1 brings additional flexibility to the active site in the ES complex. This is pronounced in wider distributions of key interatomic distances, such as the distance of the nucleophilic attack, coordination bond lengths, and covalent bond lengths in the substrate. Substrate activation, quantified by Fukui electrophilicity index of the carbonyl carbon atom of the substrate, is also sensitive to the active site flexibility. In the tighter and more rigid L1 enzyme-substrate complex, the substrate is activated more efficiently. In the NDM-1 containing system, only one third of the states are activated to the same extent. Other fractions demonstrate lower substrate activation. Efficiency of the substrate activation and rigidity of the ES complex influence the following chemical reaction. In the more rigid L1-containing system, the reaction barrier of the first step of the reaction is lower, and the first intermediate is more stabilized compared to the NDM-1 containing system.

Keywords: metallo-β-lactamase; QM/MM molecular dynamics; imipenem; active site flexibility; substrate activation; machine learning

1. Introduction

Enzymes that belong to the same family share similar composition of active sites. Still, differences in the overall structure may affect their flexibility and particular conformations of the active site residues. The simplest model of the enzyme-substrate interaction is the “lock and key”, which assumes that the enzyme is quite rigid and compatible with a certain substrate. The idea of the enzyme flexibility was introduced later by Koshland in his “induced fit” model [1]. Accumulation of more than 60,000 crystal structures of different enzymes in the protein data bank makes it possible to perform big data analysis and explicitly demonstrate flexibility of amino acid residues [2–6]. These studies are based on the analysis of crystal structures and are related to the stable states like apo and holo forms, enzyme-inhibitor complexes, etc., but not metastable reactive species.

These studies encouraged us to deepen the understanding of the role of active site flexibility, focusing on the reactive states. To do this, we selected two bacterial enzymes related to the antibiotic resistance, metallo-β-lactamases L1 (a B3 metallo-β-lactamase from Stenotrophomonas (Xanthomonas) maltophilia) and NDM-1 (New Delhi metallo-β-lactamase), belonging to the same B class [7]. The reaction mechanism of the imipenem antibiotic hydrolysis by these two enzymes has already been studied [8]. The consequence of
the elementary steps is the same for both enzymes; still, the energy barrier at the first step is lower for the L1-containing system. Also, the first intermediate is stabilized relative to the enzyme-substrate complex (ES) in the L1-imipenem complex and destabilized in the NDM-1. In both enzymes, the L10 loop that is conservative for metallo-β-lactamases covers the substrate and fixes it in the active site (Figure 1, [9]). In the L1, this loop contains a proline residue right above the substrate that brings rigidity to the construct (Figure 1A). Conversely, in the NDM-1, this loop is more flexible, as it carries glycine residue at the same position, and the pyrroline ring of imipenem is more exposed to the solution in the NDM-1 (Figure 1B). Both enzymes contain two zinc cations, Zn12+ and Zn22+, bridged by a catalytic hydroxide ion, OwH− (Figure 1). Ligands of Zn12+ cations are the same in both enzymes: three histidine residues, catalytic hydroxide anion, and a carbonyl group of the substrate. The coordination surrounding the Zn22+ cation differs only in one ligand: a histidine residue in the L1 and a cysteine residue in the NDM-1. Others are the same; those are a histidine residue, a catalytic OwH−, a catalytic aspartate residue, and a nitrogen atom of the substrate (Figure 1).

Figure 1. Active sites of the L1 (A) and the NDM-1 (B) with imipenem. The fragment of the L10 loop is shown in red for the L1 and in dark blue for the NDM-1. Coordination spheres of zinc cations and side chains of amino acid residues forming hydrogen bonds with the carboxyl group of imipenem are depicted with thin sticks. The N–C=O fragment of the β-lactam ring that participates in chemical reaction is shown in thick sticks. Here, carbon is green, oxygen is red, nitrogen is blue, hydrogen is white, sulfur is yellow, and zinc is dark gray.

In enzyme-substrate complexes of both enzymes, the oxyanion hole is formed by the Zn12+ cation, and its interactions with the carbonyl oxygen atom polarize the C=O bond of the substrate (Figure 2). Additionally, the neighboring C-N bond is polarized due to the interactions between its nitrogen atom and a Zn22+ cation (Figure 2). These interactions prepare or activate the carbonyl carbon atom of the β-lactam ring for nucleophilic attack that initiates the reaction. Polarization of both C=O and C-N bonds results in the increase of the negative partial charges on heteroatoms (δ−) and increase of the partial positive charge (δ+) on the carbon atom. Thus, the carbonyl carbon becomes more electrophilic, which facilitates nucleophilic attack by the catalytic OwH−. The nucleophilic particle OwH− attacks a carbonyl carbon atom, leading to the formation of the tetrahedral intermediate. Substrate activation can be quantified in different ways, including simple geometry criteria and more complicated electron density-based descriptors. Interatomic distances in the
oxyanion hole and distances of the nucleophilic attack can be utilized as a measure of interactions of the carbonyl group of the substrate with the active site of the enzyme. Among electron density-based criteria, here, we focus on Laplacian of electron density [10–14] and Fukui atomic indices [15–17] (for their applications see, for example, refs [13,18–22]). Analysis of Laplacian of electron density, \( \nabla^2 \rho(r) \), allows identification of the areas of the local electronic charge concentration with \( \nabla^2 \rho(r) < 0 \) (nucleophilic sites) and electronic charge depletion areas with \( \nabla^2 \rho(r) > 0 \) (electrophilic sites) in molecular systems. Analysis of \( \nabla^2 \rho(r) \) 2D maps in the plane formed by the nucleophilic atom of the catalytic particle and the carbonyl group of the substrate provides easily visible images with a binary classification as reactive and nonreactive species. Fukui atomic indices are utilized for quantitative estimates of electrophilic and nucleophilic site. Here we use Fukui electrophilicity index to quantify the extent of activation of the carbonyl carbon atom of the substrate.

Figure 2. Scheme of the activation of the carbapenem by metallo-β-lactamases L1 and NDM-1. Black arrows on the left part of the figure shows the direction of charge concentration shift; on the right side, the nucleophilic attack of the activated carbonyl carbon by \( \text{OwH}^- \) is shown. Black dashed arrow shows the nucleophilic attack on the non-activated carbonyl carbon, blue dashed lines show coordination bonds of zinc ions with the substrate in the active site. Red molecular fragment polarizes in the active site, thus activating the carbonyl carbon atom.

In this study, we compare the flexibility of two enzyme-substrate complexes and reveal how the flexibility affect the energetic of the subsequent chemical reaction. We rely on both geometry criteria and electron density-based features. We calculate interatomic distances in the enzymatic active site and obtain electron density-based features that are responsible for reactivity in sets of QM/MM MD frames. We utilize Fukui electrophilicity index and Laplacian of electron density as descriptors and analyze how geometry criteria are interrelated with the substrate activation quantified by the Fukui electrophilicity index.

2. Results and Discussion

2.1. Dynamic Behavior of the ES Complexes: Geometry Parameters

Dynamic behavior of the ES complexes can shed light on the subsequent chemical reaction. The most straightforward approach is to analyze key interatomic distances and compare them among similar systems. Here, we compare dynamic behavior of the ES complexes of the L1 (L1-ES) and NDM-1 (NDM-1-ES) metallo-β-lactamases with the imipenem substrate.

Interactions between the oxyanion hole and the carbonyl oxygen atom play a decisive role in the preparation of the substrate for a nucleophilic attack [23–25]. In both ES complexes, the oxyanion hole is represented by a Zn1\(^{2+} \) cation that forms a coordination bond
with the carbonyl oxygen atom of the substrate (Figures 2 and 3D, Table 1). Also, there is another interaction that should increase electrophilicity of the carbonyl carbon atom due to the polarization of the C-N bond of the substrate: a coordination bond between Zn$^{2+}$ and the nitrogen atom of the β-lactam ring (Figures 2 and 3E, Table 1). In the L1-ES, d(Zn$^{2+}$…O) and d(Zn$^{2+}$…N), distributions are narrow, the distances vary within 0.21 Å and 0.14 Å, respectively. The average values are similar, being 2.39 ± 0.04 Å for the Zn$^{2+}$…O distance and 2.36 ± 0.02 Å for the Zn$^{2+}$…N distance. These values correspond to coordination bonds. In the NDM-1-ES, the distributions are much wider, the bond lengths vary within 1.32 Å and 0.9 Å for d(Zn$^{2+}$…O) and d(Zn$^{2+}$…N), respectively. Moreover, the Zn$^{2+}$…O distances reach the values of 4 Å. Practically, no coordination bond exists between the zinc cation and the oxygen atom of the substrate, and the only influence is due to the large positive charge of the metal cation, that is, the C=O bond polarization occurs due to the electrostatic field of the Zn$^{2+}$. The Zn$^{2+}$…N distance is shorter than Zn$^{2+}$…O distance; their mean values are 2.52 ± 0.18 Å and 3.18 ± 0.26 Å, respectively. We assume that the Zn$^{2+}$…N coordination bond may partially compensate the absence of the Zn$^{2+}$…O coordination bond. The C-N bond polarization should also increase the extent of the charge deconcentration on the carbonyl carbon atom in the NDM-1, thus making it an electrophile prepared for the reaction.

**Figure 3.** Distributions of interatomic distances in the L1-ES (red) and the NDM-1-ES (blue): (A) distance of the nucleophilic attack, d(C…Ow); (B) C=O bond length in the carbonyl fragment of the β-lactam ring, d(C=O); (C) C-N bond length of the β-lactam ring, d(C-N); (D) coordination bond length between the zinc cation and the oxygen atom of the carbonyl group of the substrate, d(Zn$^{2+}$…O) (E) coordination bond length between the zinc cation and the nitrogen atom of the β-lactam ring, d(Zn$^{2+}$…N) (F) difference between Zn$^{2+}$…O and Zn$^{2+}$…N coordination bond lengths. All values are from QM/MM MD trajectories of the ES complexes.
Table 1. Key interatomic distances obtained in the QM/MM MD simulations (with “MD” superscript) and at the minima on the potential energy surfaces corresponding to the ES complexes [8] (with “PES” superscript) for the L1-ES and the NDM-1-ES. In the rows with the QM/MM MD results the first line is for the range of corresponding distance and the second for the average value and standard deviation. The third line in the NDM-1MD row corresponds to the decomposition of the distance distributions into three states (weight of each state is in parenthesis).

| System  | d(C…)Ow | d(Zn12+…O) | d(Zn22+…N) | d(C=O)   | d(C-N)   |
|---------|---------|-------------|-------------|----------|----------|
| L1MD    | 2.62–2.76 | 2.29–2.50   | 2.29–2.43   | 1.201–1.225 | 1.381–1.421 |
|         | 2.68 ± 0.03 | 2.39 ± 0.04 | 2.36 ± 0.02 | 1.212 ± 0.003 | 1.402 ± 0.007 |
|         | 2.88 ± 0.15 | 3.23 ± 0.25 | 2.55 ± 0.18 | 1.210 ± 0.021 | 1.417 ± 0.036 |

| L1PES   | 2.57–3.30 | 2.65–3.94   | 2.23–3.13   | 1.150–1.270 | 1.324–1.564 |
|         | 2.88 ± 0.04 | 3.25 ± 0.02 | 2.57 ± 0.05 | 1.197 ± 0.0001 | 1.3847 ± 0.0007 |

| NDM-1MD | 2.71 ± 0.004 | 2.99 ± 0.02  | 2.391 ± 0.005 | 1.1970 ± 0.0001 | 1.3847 ± 0.0007 |
|         | (0.29)      | (0.30)      | (0.40)       | (0.28)       | (0.38)       |
|         | 2.87 ± 0.004 | 3.21 ± 0.02  | 2.579 ± 0.006 | 1.2094 ± 0.0001 | 1.4260 ± 0.0004 |
|         | (0.42)      | (0.39)      | (0.40)       | (0.42)       | (0.43)       |
|         | 3.05 ± 0.007 | 3.49 ± 0.04  | 2.81 ± 0.02  | 1.2318 ± 0.0001 | 1.4628 ± 0.0009 |
|         | (0.29)      | (0.30)      | (0.21)       | (0.30)       | (0.19)       |

| NDM-1PES | 2.74 | 3.59 | 2.24 | 1.200 | 1.455 |

The distance of the nucleophilic attack between the carbonyl carbon atom of the substrate and the catalytic OwH is another important feature of the ES complex. In the L1-ES, the average value of the nucleophilic attack distance is 2.67 ± 0.03 Å, the distribution is narrow, values vary in a range of 2.62–2.76 Å. For the NDM-1, the distribution is much wider, the range is 2.57–3.30 Å, the average and the standard deviation are 2.85 ± 0.15 Å. Mean values for both enzymes are close, but the shape of the distributions is considerably different (Figure 3A). In the L1-ES, distance of the nucleophilic attack varies within only 0.14 Å; for the NDM-1-ES, it increases to 0.73 Å.

Covalent bonds that are polarized in the active sites are also considered. Those are C=O bond and C-N bond of the β-lactam ring. Similarly to the distances distributions corresponding to the noncovalent interaction, we observe narrow distributions for the L1-ES and wide distributions for the NDM-1-ES for both C=O and C-N bonds (Figure 3B,C, Table 1). Despite the different shapes of the carbonyl bond length distributions, average values are the same for both enzymes. Those are 1.212 ± 0.003 Å and 1.210 ± 0.021 Å for the L1-ES and the NDM-1-ES, respectively. The maximum of the C=N bond length distribution is shifted towards larger values in the NDM-1-ES compared with the L1-ES. The average values here are 1.402 ± 0.007 Å and 1.418 ± 0.036 Å for the L1-ES and the NDM-1-ES, respectively. The difference of 0.16 Å in the average values is significant for a covalent bond in a cyclic fragment of a molecule. The origin of this difference might lie in the cooperative effect of coordination bonds with zinc cations (Figure 3F). To demonstrate this, we calculate the difference between d(Zn12+…O) and d(Zn22+…N). If it is close to zero, both coordination bonds have the same lengths and thus contribute similarly to the polarization of the C-N and C=O bonds. In the NDM-1-ES, the distribution of the difference of coordination bond lengths is considerably shifted from zero. Thus, in the NDM-1, the effect of the Zn22+…N coordination bond is more pronounced compared with the L1-ES. Indeed, the distance distribution of the C-N bond is wider and shifted to larger values; C-N bond lengths are observed for systems with shorter Zn22+…N (Figure 4C).
Figure 4. (A) One-state and three-state models of distribution of the nucleophilic attack distance, C...Ow, in the NDM-1-ES. Magenta line corresponds to the unimodal normal distribution, and navy solid line is a sum of three normal distributions (navy dashed lines) with the corresponding weights. Scatter plots of interatomic distances in the MD trajectory of the NDM-1-ES; (B) distance of the nucleophilic attack, d(C...Ow) vs Zn$^{2+}$...O (green) or Zn$^{2+}$...N (blue) coordination bond, d; (C) C–N bond vs Zn$^{2+}$...N bond.

For all considered interatomic distances, the same tendency is observed. In the L1-ES, distance distributions are narrower than for the NDM-1-ES. We suppose that this is due to the different organization of the active site region. In the L1, the substrate is additionally fixed by hydrophobic interactions with the Pro226 residue of the loop 10 (Figure 1). These interactions restrict significant movements of the imipenem in the active site of the L1. Conversely, in the NDM-1, Gly219 residue is located above the substrate (at the same position as Pro226 in the L1), and the loop 10 is located further from the substrate (Figure 1). Thus, in the NDM-1, the imipenem is generally more flexible compared with the L1 that is also pronounced in the distributions of key interatomic distances.

The detailed analysis of distance distributions for the NDM-1-ES (Figure 3) reveals that they are better described as a sum of normal distributions rather than one. Thus, structural flexibility of the active site leads to the heterogeneity of the ES complexes that is seen in the distributions of geometry parameters. Figure 4A and Table 1 demonstrate that there are three populations in the NDM-1 ES with respect to the C...Ow distance with similar weights. Similarly, multimodal distributions of other interatomic distances discussed above are observed. We found no good correlations between considered distances. The poor one was found between C-N bond lengths and the length of coordination bond that polarizes this bond (Figure 4C). Also, we found that, generally, structures with shorter distances of the nucleophilic attack have smaller coordination bond lengths (Figure 4B). It means that the active site is fluctuating between tighter states that are more reactive and looser conformations with less activated substrate.

2.2. Dynamic Behavior of the ES Complexes: Electron Density Features

Electron density descriptors, such as Fukui electrophilicity index and Laplacian of electron density, can quantify the activation of the substrate by the enzyme in its active site. We calculated Fukui electrophilicity indices at 500 MD frames for each system (Figure 5). Higher electrophilicity index of the carbonyl carbon atom of the imipenem is an indication of the more electrophilic atom, that is, such a carbon atom is more prepared for a nucleophilic attack. Fukui electrophilicity index distribution for the imipenem carbonyl carbon atom is narrower for the L1-ES (0.083–0.145 a.u.) than for the NDM-1-ES (0.002–0.181 a.u.), similarly to the distributions of geometry parameters discussed above (Figure 5A). Distributions are characterized by the following average values and standard deviations: 0.124 ± 0.009 a.u. for the L1 and 0.086 ± 0.036 a.u. for the NDM-1. The standard
deviation of the Fukui electrophilicity index is much larger for the NDM-1, and the distribution is better described by three states. Indeed, we found three populations with the distributions centered at 0.047 a.u., 0.081 a.u., and 0.123 a.u., with the corresponding weights of 0.29, 0.38 and 0.33. The most reactive population of the NDM-1 has similar electrophilicity index distribution as the distribution of the L1 values.

Figure 5. (A) Distributions of Fukui electrophilicity indices, \( f^+ \), on the carbonyl carbon atom of the imipenem in the ES complexes with the L1(red) and NDM-1(navy). (B,C) Representative Laplacian of the electron density maps in the plane of the carbonyl group of the substrate (C and O atoms) and an oxygen atom, Ow, of the catalytic O=H\(^-\) discriminating, (B) reactive, and (C) nonreactive ES complexes. Contour lines for the Laplacian of electron density maps are \( \pm (2; 4; 8) \times 10^n \) a.u., \(-2 \leq n \leq 1\), blue dashed contour lines indicate the electron density depletion areas (\( \nabla^2 \rho(r) > 0 \)), red solid lines identify the electron density concentration (\( \nabla^2 \rho(r) < 0 \)), and green solid line corresponds to \( \nabla^2 \rho(r) = 0 \). The area with \( \nabla^2 \rho(r) < 0 \) is colored in light green.

Laplacian of electron density, \( \nabla^2 \rho(r) \), maps in the plane of carbonyl group of the substrate (C and O atoms) and the nucleophilic atom, Ow, allow binary classification of species as reactive and nonreactive. We calculated these 2D maps and classified each structure as either reactive or nonreactive for a set of 500 frames from QM/MM MD trajectory of the ES complexes of both enzymes. Reactive species (Figure 5B) are characterized by the electron density deconcentration (\( \nabla^2 \rho(r) > 0 \)) in the carbonyl carbon atom region between the C and Ow nuclei. On the map, this is depicted as the area colored in white with blue dashed lines. For nonreactive species, the carbonyl carbon atom is enveloped by the electron density concentration region (\( \nabla^2 \rho(r) < 0 \), colored in light green with red solid lines, Figure 5C). We found that all considered frames from the QM/MM MD trajectory of the L1-ES complex are reactive (100%), and for the NDM-1-ES the major fraction is reactive (90%).

Results of electron density analysis are interrelated with key interatomic distances in the active sites of two considered complexes. As noted in 3.1, the substrate is less flexible in the active site of the L1 compared with the NDM-1. For the L1-ES, narrow distributions of nucleophilic attack distances, coordination bonds, and bonds in the \( \beta \)-lactam ring are observed. In the NDM-1-ES, imipenem is more exposed to the solution and, generally, more flexible, that is, pronounced in considerably broader distributions of considered distances. The distributions of the cooperative action of coordination bonds in Figure 3F show that in the L1 both coordination bonds are strong, while the Zn\(^{2+}\)...N coordination bond dominates in the NDM-1. Hence, all these geometry features of the active site led to the different extent of substrate activation in the active site and different heterogeneity of states. In the L1-ES, 100% of states are reactive, while 90% are reactive in the NDM-1-ES. Fukui electrophilicity indices of the carbonyl carbon atoms are larger in the L1-ES compared with the NDM-1-ES. Also, for L1, we observe a unimodal distribution for both geometry and electron density features, whereas in the NDM-1 we find heterogeneity that can be described with at least 3 states.
2.3. Interrelation between Geometry and Electron Density Features in the ES

Multiple variable linear regression analysis and machine learning random forest algorithm were utilized to study the relationship between geometry features and Fukui electrophilicity index of the carbonyl carbon atom in the ES state in the NDM-1. We selected the same geometry criteria as the ones shown in Figure 3. We obtained a multivariable linear regression with the MAE 0.022 a.u. (Figure 6A). This error is large, still, it allows to discriminate qualitatively both more and less reactive species. This is useful, as calculation of Fukui indices require additional computations in doublet state with an additional electron in the system. Still, herein, we are mostly interested in understanding the nature of the substrate activation, that is, which geometry parameters are more important. To do this, we utilized machine learning algorithms. Random forest approach resulted in a much better prediction quality of Fukui indices (Figure 6B), with the MAE 0.008 a.u. Also, it allowed us to explicitly estimate which geometry parameters are important (Figure 6C). We found that no geometry parameter has an importance exceeding 0.25; both distance of nucleophilic attack and polarization of the N-C=O fragment are important for the substrate activation.

Figure 6. Calculated vs predicted Fukui indices for the NDM-1—imipenem ES complexes. (A) Predictions made from multivariable linear regression. (B) Results for the validation set in the random forest machine learning algorithm. (C) Importance of geometry parameters suggested by machine learning.

2.4. The First Reaction Step

The first step of the reaction is the nucleophilic attack of the hydroxide anion (O−H\(^{-}\)) on the carbonyl carbon atom of the \(\beta\)-lactam ring (Figure 7A). Here, we perform a set of QM/MM MD simulations with harmonic potentials centered at different values of the reaction coordinate, d(C…O\(_{\text{w}}\)). This allows us to obtain Gibbs energy profiles for the first step for both systems and compare them. We are interested in the comparison of the Gibbs energy profiles obtained for the NDM-1 with the L1-containing systems. The reaction coordinate at the first step in the umbrella sampling simulations is the distance of the nucleophilic attack, d(C…O\(_{\text{w}}\)). Energy barriers at the first step are quite low, being 9 kcal/mol for the L1-containing system and 11 kcal/mol for the NDM-1. The difference between the energy barriers of the first step of the reaction in the L1 and the NDM-1 is about 2 kcal/mol, while the intermediate, Int, is 1.8 kcal/mol stabilized relative to the ES in the NDM-1 and almost 5 kcal/mol lower in energy in the L1 (Figure 7B). Thus, the first step of the reaction occurs more efficiently in the L1-imipenem complex. This is in line with the analysis of the ES complex: the carbonyl carbon atom of the substrate is more electrophilic due to the tighter and less flexible binding of the imipenem in the L1 active site. These results are in line with the data obtained on potential energy surface [8]. The energy barrier is lower, and the intermediate stabilization energy is larger in the L1 compared with the NDM-1.
3. Methods

The enzyme-substrate complex of the L1 and imipenem was obtained from the crystal structure PDB ID: 2AIO [26], and the enzyme-substrate complex of the NDM-1 and imipenem was obtained from the crystal structure PDB ID: 5YPK [27]. In the L1-containing system, hydrolyzed moxalactam was substituted with imipenem by superimposition of atoms of the hydrolyzed β-lactam ring and neighboring carboxylate group of moxalactam with the β-lactam ring and carboxylate group of imipenem. The complex with the NDM-1 imipenem was restored to its non-hydrolyzed form. Protonation states of acidic and basic amino acids were assigned to model neutral pH as follows. Carboxyl groups of aspartate and glutamate residues were negatively charged; side chains of lysine and arginine residues were in the protonated states; histidine residues were neutral and the protonation of either δ or ε nitrogen atom was chosen depending on the local environment. In the L1, His94, His118, and His155 were Nε-protonated, others were Nδ-protonated. In the NDM-1, all histidine residues, except for His122, were Nδ-protonated. These systems were solvated in rectangular water boxes, so that the distance from the protein surface to the border of the cell exceeded 12 Å. These systems were neutralized by adding sodium ions. Systems were preliminary equilibrated during 10 ns in classical molecular dynamics (MD) simulations performed at T = 300 K and p = 1 atm with the 1 fs integration time step using CHARMM36 [28] force field for the protein, TIP3P [29] for water molecules, CGenFF [30] for the imipenem, and force field parameters for zinc cations from ref [31]. Root-mean-square deviation (RMSD) values along the obtained trajectories showed that this length of the trajectories is enough for the complete relaxation of the system. Then the representative frames from the last 2 ns of the MD runs were used to prepare models for subsequent QM/MM MD simulations.

The QM subsystem included the imipenem, two Zn2+ cations, and side chains of amino acid residues that form coordination bonds with them, catalytic hydroxide anion, and side chains that form hydrogen bonds with the imipenem (Ser221 and Tyr32 in the L1; Lys211, Leu218-Asn220, in the NDM-1). All QM/MM MD calculations were performed at the Kohn-Sham DFT level with the PBE0 functional [32] and empirical dispersion correction D3 [33], all electron 6-31G** basis set was utilized for all atom except for zinc, which was described with the pseudopotentials LANL2DZ [34]. QM/MM MD calculations were performed in the combination of TeraChem and NAMD programs [35]. The TeraChem program [36] was used to calculate forces in the QM region and NAMD in the MM part, as well as molecular dynamics steps [37]. QM/MM MD simulations were performed in the NPT ensemble at T = 300 K and p = 1 atm with the 1 fs integration time step. The cutoff distance for point charges of the MM subsystem contributing to the QM Hamiltonian was 12 Å. The VMD software [38] was used for preparation of the model systems and subsequent analysis of MD trajectories. Gibbs energy profiles for the first step,
including nucleophilic attack and formation of the tetrahedral intermediate, were calculated using the umbrella sampling approach [39]. The sets of 5–10 ps runs were performed with harmonic potentials centered at different values of the reaction coordinate, which was the distance of nucleophilic attack. The force constant of the harmonic potential \(\frac{1}{2}K(\xi - \xi_0)^2\) was usually set to 40 kcal/mol/Å², and additional trajectories with \(K = 80–320\) kcal/mol/Å² in transition state regions were calculated. Harmonic potentials were centered every 0.2 Å along the reaction coordinate. Umbrella integration (UI) and weighted histogram analysis method (WHAM) were utilized to reconstruct energy profiles of the first reaction steps.

Analysis of dynamic behavior of the ES complexes was performed as follows. We selected sets of 500 frames for each model system from the last 5 ps QM/MM MD simulations of L1 and NMD-1 containing systems in the ES configuration. We estimated reactivity of imipenem using Laplacian of electron density maps calculated in the plane of carbonyl group of the substrate (C and O atoms) and the nucleophilic atom, O_w, and Fukui electrophilicity index, \(f^+\), on the carbonyl carbon atom. Atomic index of electrophilicity, \(f^+\), is evaluated as a difference between Hirshfeld charges [40], calculated for the model system with N electrons (a system under configuration), and a system with the same geometry configuration, but with an additional electron. To do this, we performed additional single point calculations in the doublet state using unrestricted Kohn-Sham method with the same functional and basis set as calculations discussed above. Electron density analysis was performed in the Multiwfn program package [41].

Machine learning was performed for the datasets composed of calculated Fukui electrophilicity indices and key interatomic parameters at the corresponding MD frames. Random forest approach implemented in the Scikit-learn [42] was utilized to obtain a model to predict Fukui electrophilicity indices on the carbonyl carbon atom of imipenem from geometry parameters and to estimate importance of these parameters in value estimates.

4. Conclusions

We performed molecular dynamic simulations with QM/MM potentials to compare structural flexibility of active site in the ES complex of two metallo-\(\beta\)-lactamases, the L1 and the NDM-1, with the imipenem antibiotic. Active sites in these complexes are covered by L10 loop that is more rigid in the L1 due to the presence of the proline residue. Analysis of key interatomic distances in the ES complex demonstrates a general trend: distributions of all distances are narrower in the L1-ES compared with the NDM-1-ES. Structural heterogeneity of ES complexes in the NDM-1 is also pronounced in the multimodal description of distance distributions. Distributions of the Fukui electrophilicity indices share the same tendencies: a narrow unimodal distribution in the L1, and a wide distribution composed of at least three populations with comparable weights in the NDM-1. Machine learning reveals that all geometry parameters, including distance of the nucleophilic attack, coordination bond lengths between the substrate and zinc cations, and C=O and C–N covalent bond lengths in the active site have similar importance for the Fukui index estimates. Gibbs energy profiles calculated for the hydrolysis of the imipenem in the active sites of the L1 and NDM-1 shares the same features with the profiles on the potential energy surface [8]. Formation of tighter and more rigid active site in the L1 leads to the lower energy barrier and larger extent of the intermediate stabilization relative to the ES complex.

Author Contributions: Conceptualization, M.G.K.; methodology, M.G.K., A.V.K.; data curation, M.G.K.; writing—original draft preparation, M.G.K., A.V.K.; validation, M.G.K., A.V.K.; visualization, M.G.K., A.V.K.; supervision, M.G.K.; project administration, M.G.K.; funding acquisition, M.G.K. All authors have read and agreed to the published version of the manuscript.

Funding: QM/MM and QM/MM MD simulations and electron density analysis were supported by Russian Science Foundation (project №18-74-10056). Machine learning study was supported by the
Interdisciplinary Scientific and Educational School of Moscow University «Brain, Cognitive Systems, Artificial Intelligence» (M.G.K.).

Data Availability Statement: Not applicable.

Acknowledgments: The research is carried out using the equipment of the shared research facilities of HPC computing resources at Lomonosov Moscow State University and Joint Supercomputer Center of the Russian Academy of Sciences.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Koshland, D.E. The Key–Lock Theory and the Induced Fit Theory. Angew. Chemie Int. Ed. English 1995, 33, 2375–2378. https://doi.org/10.1002/anie.199423751.

2. Weng, Y.-Z.; Chang, D.T.-H.; Huang, Y.-F.; Lin, C.-W. A study on the flexibility of enzyme active sites. BMC Bioinform. 2011, 12, S32. https://doi.org/10.1186/1471-2105-12-S1-S32.

3. Gutteridge, A.; Thornton, J. Conformational Changes Observed in Enzyme Crystal Structures upon Substrate Binding. J. Mol. Biol. 2005, 346, 21–28. https://doi.org/10.1016/j.jmb.2004.11.013.

4. Gutteridge, A.; Thornton, J. Conformational change in substrate binding, catalysis and product release: An open and shut case? FEBS Lett. 2004, 567, 67–73. https://doi.org/10.1016/j.febslet.2004.03.067.

5. Riziotis, I.G.; Ribeiro, A.J.M.; Borkakoti, N.; Thornton, J.M. Conformational Variation in Enzyme Catalysis: A Structural Study on Catalytic Residues. J. Mol. Biol. 2022, 434, 167517. https://doi.org/10.1016/j.jmb.2022.167517.

6. Demetrius, L. Role of Enzyme–substrate Flexibility in Catalytic Activity: An Evolutionary Perspective. J. Theor. Biol. 1998, 194, 175–194. https://doi.org/10.1006/jtbi.1998.0748.

7. Bebrone, C. Metallo-β-lactamases (classification, activity, genetic organization, structure, zinc coordination) and their superfamily. Biochem. Pharmacol. 2007, 74, 1686–1701. https://doi.org/10.1016/j.bcp.2007.05.021.

8. Krivitskaya, A.V.; Khrenova, M.G. Interplay between the Enamine and Imine Forms of the Hydrolyzed Imipenem in the Active Sites of Metallo-β-lactamases and in Water Solution. J. Chem. Inf. Model. 2022. https://doi.org/10.1021/acs.jcim.2c00539.

9. Bahr, G.; González, L.J.; Vila, A.J. Metallo-β-lactamases in the Age of Multidrug Resistance: From Structure and Mechanism to Evolution, Dissemination, and Inhibitor Design. Chem. Rev. 2021, 121, 7957–8094. https://doi.org/10.1021/acs.chemrev.1c00138.

10. Matta, C.F.; Bader, R.F.W. Atoms-in-molecules study of the genetically encoded amino acids. III. Bond and atomic properties and their correlations with experiment including mutation-induced changes in protein stability and genetic coding. Proteins Struct. Funct. Bioinform. 2003, 52, 360–399. https://doi.org/10.1002/prot.10414.

11. Bader, R.F.W.; Popelier, P.L.A.; Chang, C. Similarity and complementarity in chemistry. J. Mol. Struct. Theorchem. 1992, 255, 145–171. https://doi.org/10.1016/0166-1280(92)89008-9.

12. Khrenova, M.G.; Kulakova, A.M.; Nemukhin, A.V. Proof of concept for poor inhibitor binding and efficient formation of covalent adducts of KRAS G12C and ARS compounds. Org. Biomol. Chem. 2020, 18, 3069–3081. https://doi.org/10.1039/D0OB00071J.

13. Carroll, M.T.; Cheeseman, J.R.; Osman, R.; Weinstein, H. Nucleophilic addition to activated double bonds: Predictions of reactivity from the Laplacian of the charge density. J. Phys. Chem. 1989, 93, 5120–5123. https://doi.org/10.1021/j100350a019.

14. Shi, Z.; Boyd, R.J. Stereoelectrochemistry of nucleophilic addition to substituted cyclohexanones: A structure and charge density study. J. Am. Chem. Soc. 1993, 115, 9614–9619. https://doi.org/10.1021/ja00074a029.

15. Liu, S.; Rong, C.; Lu, T. Information Conservation Principle Determines Electrophilicity, Nucleophilicity, and Regioselectivity. J. Phys. Chem. A 2014, 118, 3698–3704. https://doi.org/10.1021/jp5032702.

16. Roy, R.K. Stockholders Charge Partitioning Technique. A Reliable Electron Population Analysis Scheme to Predict Intramolecular Reactivity Sequence. J. Phys. Chem. A 2003, 107, 10428–10434. https://doi.org/10.1021/jp035848z.

17. Oláh, J.; Van Alsene, C.; Sannigrahi, A.B. Condensed Fukui Functions Derived from Stockholder Charges: Assessment of Their Performance as Local Reactivity Descriptors. J. Phys. Chem. A 2002, 106, 3885–3890. https://doi.org/10.1021/jp014039h.

18. Khrenova, M.G.; Tsirelson, V.G.; Nemukhin, A.V. Dynamical properties of enzyme–substrate complexes disclose substrate specificity of the SARS-CoV-2 main protease as characterized by the electron density descriptors. Phys. Chem. Chem. Phys. 2020, 22, 19069–19079. https://doi.org/10.1039/D0CP03560B.

19. Khrenova, M.G.; Krivitskaya, A.V.; Tsirelson, V.G. The QM/MM-QTAIM approach reveals the nature of the different reactivity of cephalosporins in the active site of L1 metallo-β-lactamase. New J. Chem. 2019, 43, 7329–7338. https://doi.org/10.1039/C9NJ00254E.

20. Krivitskaya, A.V.; Khrenova, M.G. Boronic Acids as Prospective Inhibitors of Metallo-β-Lactamases: Efficient Chemical Reaction in the Enzymatic Active Site Revealed by Molecular Modeling. Molecules 2021, 26, 2026. https://doi.org/10.3390/ijms23021001.
Molecules 2022, 27, 7031

22. Milanović, Ž.B.; Dimić, D.S.; Avdović, E.H.; Milenković, D.A.; Marković, J.D.; Klisurić, O.R.; Trifunović, S.R.; Marković, Z.S. Synthesis and comprehensive spectroscopic (X-ray, NMR, FTIR, UV–Vis), quantum chemical and molecular docking investigation of 3-acyl-4-hydroxy-2-oxo-2H-chromen-7-yl acetate. J. Mol. Struct. 2021, 1225, 129256. https://doi.org/10.1016/j.molstruc.2020.129256.

23. Kamerlin, S.C.L.; Chu, Z.T.; Warshel, A. On Catalytic Preorganization in Oxyanion Holes: Highlighting the Problems with the Gas-Phase Modeling of Oxyanion Holes and Illustrating the Need for Complete Enzyme Models. J. Org. Chem. 2010, 75, 6391–6401. https://doi.org/10.1021/jo100651s.

24. Simón, L.; Goodman, J.M. Hydrogen-bond stabilization in oxyanion holes: Grand jeté to three dimensions. Org. Biomol. Chem. 2012, 10, 1905. https://doi.org/10.1039/c2ob06717j.

25. Amyes, T.L.; Richard, J.P. Specificity in Transition State Binding: The Pauling Model Revisited. Biochemistry 2013, 52, 2021–2035. https://doi.org/10.1021/bi301491r.

26. Spencer, J.; Read, J.; Sessions, R.B.; Howell, S.; Blackburn, G.M.; Gamblin, S.J. Antibiotic Recognition by Binuclear Metallo-β-Lactamases Revealed by X-ray Crystallography. J. Am. Chem. Soc. 2005, 127, 14439–14444. https://doi.org/10.1021/ja0536062.

27. Feng, H.; Liu, X.; Wang, S.; Fleming, J.; Wang, D.-C.; Liu, W. The mechanism of NDM-1-catalyzed carbapenem hydrolysis is distinct from that of penicillin or cephalosporin hydrolysis. Nat. Commun. 2017, 8, 2242. https://doi.org/10.1038/s41467-017-02339-w.

28. Best, R.B.; Zhu, X.; Shim, J.; Lopes, P.E.M.; Mittal, J.; Feig, M.; MacKerell, A.D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone φ, ψ and Side-Chain χ1 and χ2 Dihedral Angles. J. Chem. Theory Comput. 2012, 8, 3257–3273. https://doi.org/10.1021/ct300400x.

29. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 1983, 79, 926–935. https://doi.org/10.1063/1.445869.

30. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM general force field (CGenFF): A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. J. Comput. Chem. 2010, 31, 671–690. https://doi.org/10.1002/jcc.21367.

31. Stote, R.H.; Karplus, M. Zinc binding in proteins and solution: A simple but accurate nonbonded representation. Proteins Struct. Funct. Genet. 1995, 23, 12–31. https://doi.org/10.1002/prot.340230104.

32. Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. J. Chem. Phys. 1999, 110, 6158. https://doi.org/10.1063/1.478522.

33. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J. Chem. Phys. 2010, 132, 154104. https://doi.org/10.1063/1.3382344.

34. Hay, P.J.; Wadt, W.R. Ab initio effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. J. Chem. Phys. 1985, 82, 299. https://doi.org/10.1063/1.448975.

35. Melo, M.C.R.; Bernardi, R.C.; Rudack, T.; Scheurer, M.; Ripplinger, C.; Phillips, J.C.; Maia, J.D.C.; Rocha, G.B.; Ribeiro, J.V.; Stone, J.E.; et al. NAMD goes quantum: An integrative suite for hybrid simulations. Nat. Methods 2018, 15, 351–354. https://doi.org/10.1038/nmeth.4638.

36. TeraChem v 1.9, PetaChem, LLC. Available online: www.petachem.com (accessed on 19 September 2022).

37. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. J. Comput. Chem. 2005, 26, 1781–1802. https://doi.org/10.1002/jcc.20289.

38. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. J. Mol. Graph. 1996, 14, 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.

39. Kästner, J. Umbrella sampling. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2011, 1, 932–942. https://doi.org/10.1002/wcms.66.

40. Hirshfeld, F.L. Bonded-atom fragments for describing molecular charge densities. Theor. Chim. Acta 1977, 44, 129–138. https://doi.org/10.1007/BF00549096.

41. Lu, T.; Chen, F. Multitwin: A multifunctional wavefunction analyzer. J. Comput. Chem. 2012, 33, 580–592. https://doi.org/10.1002/jcc.22885.

42. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 2011, 12, 2825–2830.