

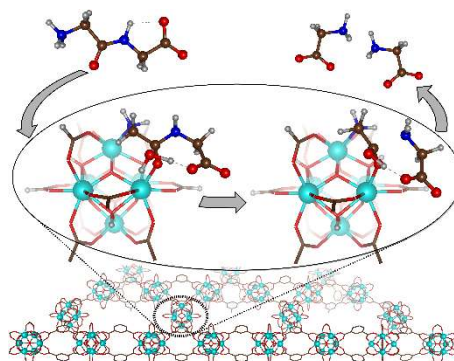
Mechanism of the highly effective peptide bond hydrolysis by MOF-808 catalyst under biologically relevant conditions

Dragan Conic,^a Kristine Pierloot,^a Tatjana N. Parac-Vogt,^a *^b Jeremy N. Harvey^{*a}

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Efficient and selective hydrolysis of inert peptide bonds is of paramount importance. MOF-808, a metal-organic framework based on Zr₆ nodes, can hydrolyze peptide bonds efficiently under biologically relevant conditions. However, the details of the catalyst structure and of the underlying catalytic reaction mechanism are challenging to establish. By means of DFT calculations we first investigate the speciation of the Zr₆ nodes and identify the nature of ligands that bind to the Zr₆O₈H_{4-x} core in aqueous conditions. The core is predicted to strongly prefer a Zr₆O₈H₄ protonation state and to be predominantly decorated by bridging formate ligands, giving Zr₆(μ₃-O)₄(μ₃-OH)₄(BTC)₂(HCOO)₆ and Zr₆(μ₃-O)₄(μ₃-OH)₄(BTC)₂(HCOO)₅(OH)(H₂O) as the most favorable structures at physiological pH. The GlyGly peptide can bind MOF in several different ways, with the preferred structure involving coordination through the terminal carboxylate analogously to the binding mode of formate ligand. The pre-reactive binding mode in which the amide carbonyl oxygen coordinates the metal core lies 7 kcal higher in free energy. The preferred reaction pathway is predicted to have two close-lying transition states, either of which could be the rate-determining step: nucleophilic attack on the amide carbon atom and C-N bond breaking, with calculated relative free energies of 31 and 32 kcal/mol, respectively. Replacement of formate by water and hydroxide at the Zr₆ node is predicted to be possible, but does not appear to play a role in the hydrolysis mechanism.



By means of DFT calculations, in this work we provide insight into the speciation of the Zr₆ nodes in MOF-808 under the aqueous reaction conditions, and reveal the detailed mechanism of its nanozymatic activity towards peptide bond hydrolysis in glycine-glycine as a model substrate

Introduction

The selective and efficient hydrolysis of peptide bonds in small peptides, oligopeptides and proteins is of huge importance in a wide variety of fields, such as protein sequencing and identification,^{1–4} food industry,⁵ medicine,⁶ cleaning industry⁷ and leather processing,⁸ to mention a few. With an estimated

half-life of up to 600 years in the absence of catalyst,⁹ the peptide bond has a remarkable stability under physiological conditions, making its hydrolysis a challenging task. Although proteolytic enzymes exhibit high catalytic activity, their self-digestion and lack of selectivity often lead to the production of very short fragments and to sample contamination, making it difficult to identify the sequence of the starting protein.¹⁰ Likewise, commonly-used chemical reagents often suffer from a variety of shortcomings, such as a requirement for harsh conditions, toxicity, and/or low yields.^{10,11} To tackle these issues and meet increasing requirements, many new materials have been synthesized and tested for hydrolysis of peptide bonds in the quest for new artificial proteases. These catalysts include transition metal and lanthanide ions, their complexes, and metal-substituted polyoxometalates,^{11–17} but despite the

^a Division of Quantum Chemistry and Physical Chemistry, Department of Chemistry, KU Leuven, Celestijnenlaan 200F, 3001 Leuven, Belgium. E-mail: jeremy.harvey@kuleuven.be

^b Laboratory of Bioinorganic Chemistry, Department of Chemistry, KU Leuven, Celestijnenlaan 200F, 3001 Leuven, Belgium

Electronic Supplementary Information (ESI) available: [Detailed computational work and Cartesian coordinates are reported.]. See DOI: 10.1039/x0xx00000x

significant progress made in improving selectivity and hydrolytic reaction rates, issues like catalyst recyclability, precipitation, product separation and purification still remain as a challenge.

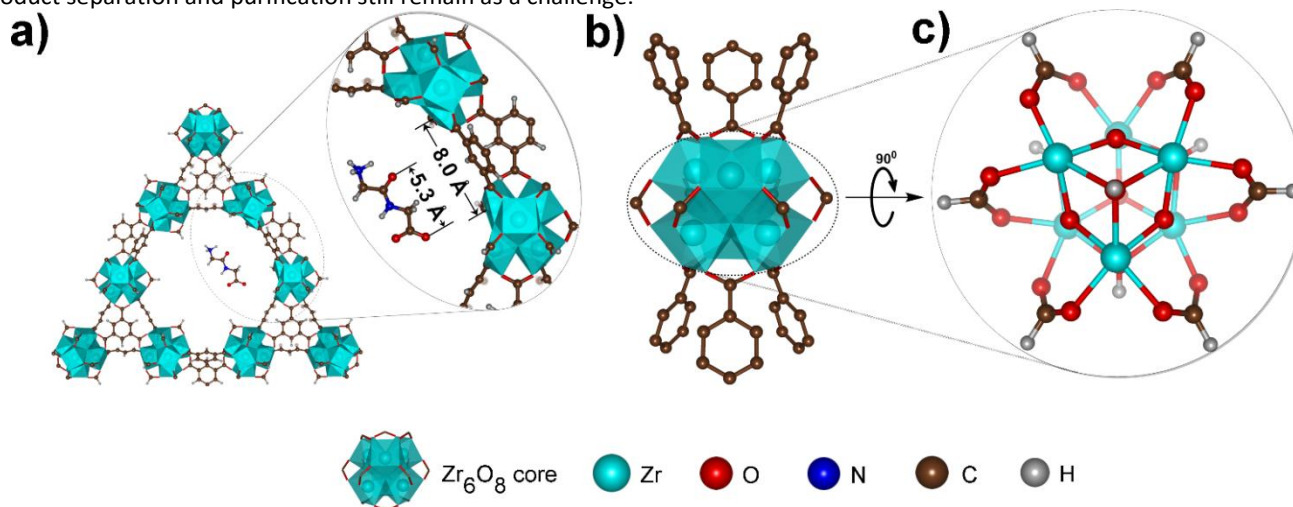


Figure 1 Polyhedral and ball-and-stick representations of MOF-808 and GlyGly dipeptide. a) GlyGly inside the MOF-808 pore. b) Single Zr_6 -core unit with its first coordination sphere (terminal carboxylic groups of the BTC linkers and hydrogen atoms are omitted for clarity). c) a dangling ligand decoration of the Zr_6 -core viewed from a base of the double cone-like shape of the linkers (linkers are omitted for clarity).

Metal-organic frameworks (MOFs) have sparked very rapidly growing research interest over the past two decades because of their structural versatility, good thermal stability, nanoporous structure and high surface area.^{18–20} As such, they show great promise for a broad range of applications, particularly in the field of catalysis.^{21–24} MOFs are a class of hybrid inorganic-organic materials, comprising metal-containing inorganic cores, designated as nodes (also known as secondary building units, or SBUs), and organic linkers connecting the nodes.²⁵ By modifying the building units, one can tailor the design in order to suit it for specific applications, for example by incorporating Lewis-acidic ions into the network for catalytic purposes. In the past few years, special attention has been directed towards the Zr_6O_8 -based MOFs because of their outstanding catalytic performance towards hydrolysis of peptide and phosphor-ester bonds in proteins, and warfare agents, respectively, as well as because of their potential in water adsorption and delivery.^{26–29} Among Zr_6O_8 -based MOFs, MOF-808 (Figure 1) has sparked considerable research interest, largely because of its superior catalytic properties in this family of MOFs.^{26,30–32} Recent studies reported superactivity of the MOF-808 towards peptide bond hydrolysis, making MOFs substantially more competitive as catalysts than previously tested.^{26,33} This was attributed to the low network connectivity, in contrast to the other Zr_6O_8 -based MOFs,³⁴ leading to the presence of a relatively high number of available coordination sites at the nodes which can be presumed to contribute to the observed catalytic activity. Moreover, MOF-808 features high thermal stability, easy product purification and good recyclability. However, the underlying mechanism of hydrolysis remains unexplored. In addition, the exact structure of the MOF-808 under physiological conditions is still to be elucidated. Although substantial research effort has been devoted to determining the nature of the loosely-bonded (dangling) ligands of the activated

as-synthesized MOF-808, a range of different structural models have been proposed in the literature.^{26,27,31,35,36} Two studies on MOF-808 by Yaghi et al. have reported two slightly different dangling ligand decorations.^{30,35} The first study proposed that the framework is described by the formula $Zr_6(\mu_3-O)(\mu_3-OH)_4(BTC)_2(HCOO)_6$ (Figure 1c), with BTC being 1,3,5-benzenetricarboxylate,³⁵ whereas the second study suggested a $Zr_6(\mu_3-O)_5(\mu_3-OH)_3(BTC)_2(HCOO)_5(H_2O)_2$ form for the system.³⁰ The difference between the two suggested structures was ascribed to slight differences in the synthetic procedures and to the possible replacement of a formate group by a molecule of the solvent used in the synthesis, DMF or water, in one of the cases. In contrast to these reports, Farha et al. argued that following the same synthetic procedure as in Yaghi et al.,³⁵ all the formate ligands might be exchanged for water molecules and hydroxyl ions by heating the material in fresh solvent.²⁷ A similar observation was made in a recent work by Chen et al., suggesting the formula $Zr_6O_4(OH)_4(BTC)_2(OH)_6(H_2O)_6$ for MOF-808.³⁶ In conclusion, despite extensive research on MOF-808 structural characterization, the nature of the dangling ligands coordinated to the Zr_6 nodes and their configuration still remain unclear.

Given the above background concerning the mechanisms of hydrolysis and the structure of the MOF, this study aims to achieve two goals: first, to shed more light on the nature and arrangement of the dangling ligands surrounding the Zr_6 node in MOF-808 under the conditions used for hydrolysis reaction, and second, to elucidate the mechanism of efficient hydrolysis of the GlyGly dipeptide catalyzed by MOF-808. For the second goal, we first investigate the structure and energetics of complexes between the dipeptide and the Zr_6 nodes, then in a second phase examine the actual hydrolytic reactivity by constructing free energy profiles for different possible reaction paths. While our earlier study contained some computational

aspects,²⁶ the structural characterization and mechanistic investigation aspects are completely new.

Computational details

Model System Choice

An initial cluster model was constructed from the reported X-ray structure of MOF-808.³⁵ The model comprised a single Zr₆-core unit, considering that the quite large distance between neighboring nodes (~8 Å for the closest zirconium atoms, Figure 1a) makes the simultaneous involvement of two nodes during dipeptide hydrolysis unlikely. Therefore, a Zr₆O₈(BTC)₆(HCOO)₆ unit was extracted from the reported structure. The BTC ligands were then replaced by BzO (benzoate, C₆H₅COO⁻) ligands, in common with previous computational studies for a range of MOFs.²⁹ As they are missing in the X-ray structure, four hydrogen atoms were added to the inorganic core, arranged in a staggered fashion, in line with the suggestion made previously by others,³⁷ to yield the neutral cluster model Zr₆O₄(OH)₄(BzO)₆(HCOO)₆. In this way, the model features both net charge neutrality and good first coordination sphere charge compensation within the Zr₆O₄(OH)₄ node. All subsequent MOF-808 model structures were derived from the 'base' Zr₆O₄(OH)₄(BzO)₆(HCOO)₆ model just described, by exchanging and/or reorganizing dangling formate ligands and the inorganic core protons. The BTC to BzO modification step was thereby performed only once, to generate the 'base' initial model, so that the same coordinates were used for the benzoate phenyl groups for all subsequent structures (see next paragraph for a discussion of the restraints placed on the phenyl groups). In this way, possible artefacts originating from this modification were expected to cancel out in calculations of the free energy changes.

Computational protocol

Geometry optimizations and frequency analysis were performed with the BP86^{38,39} functional, the GD3BJ^{40,41} Becke-Johnson damping version of Grimme's dispersion correction, and the def2-SVP⁴² all-electron basis set for all atoms except zirconium for which a valence basis set taken from the same def2-SVP family was used in combination with the corresponding MWB28 Stuttgart-Dresden effective core potential (ECP).⁴³ Initial exploratory calculations were performed without using any geometrical restraints, and showed no significant change in the positions of the phenyl groups within the benzoate models of the BTC linkers. This is in line with a previous AIMD study on a similar MOF, where only a minor twisting motion of the linkers was observed in the course of the simulation.⁴⁴ Nevertheless, in order to mimic the periodicity of the MOF-808 structure, and to avoid artefacts, the atomic positions of all eleven atoms of the C₆H₅ group atoms in all benzoate ligands were kept constant during all reported geometry optimizations.^{29,37} The calculations were accelerated by applying the Split-RI-J⁴⁵ variant of the resolution of identity (RI)^{46,47} approximation with the def2/J⁴⁸ auxiliary basis set, whereas an increased integration grid, Grid 4, was used to

integrate the pure DFT exchange-correlation term. For the mechanistic study, extra tightscf and nfinalgrid criteria were used during the geometry optimizations with the ORCA 4.1.0.^{49,50} program package so as to get reliable structures for the transition states. Subsequently, electronic energies were refined by means of single-point calculations with the B3LYP⁵¹⁻⁵³ hybrid functional complemented with the previously mentioned GD3BJ correction and a triple-zeta quality basis set, def2-TZVP.⁴² The single point calculations were additionally accelerated by means of the RIJCOSX⁵⁴ approximation, with the grid resolutions for the DFT exchange-correlation and the Hartree-Fock exchange numerical integration increased to Grid 5 and Gridx5, respectively. The effect of the reaction environment was accounted for in all calculations by means of the SMD⁵⁵ solvation model, as implemented in the ORCA software. Although the true environment is a heterogeneous medium comprised of the MOF framework and water, we treated it by using standard SMD parameters for pure water. Harmonic frequencies were computed at the same level of theory as used for optimization, BP86-D3/def2-SVP with SMD. As explained in the SI, the frequency analysis for our models typically returned several imaginary frequencies, whose cause can be traced back either to the use of constraints for the benzoate groups, or to numerical issues with the implementation of the implicit solvent model. Full details of these frequencies and the steps taken to avoid corresponding artefacts in the presented free energies are given in SI. The calculations were performed utilizing the quantum chemistry program suite ORCA 4.1.0.

Free energy calculations

The aqueous-phase free energies were calculated as a sum of individual terms presented in equation (1).

$$G^*(\text{aq}) = E_{\text{el}}^{\text{SMD-B3LYP-GD3BJ/def2-TZVP}} + G_{\text{therm}} + \Delta G^{0 \rightarrow *}$$
 (1)

The first term on the right side of the equation represents the electronic energy (with SMD solvation corrections) derived from the single point calculation. The second term comprises thermal contributions to the free energies computed using the molecular structures and vibrational frequencies, together with standard statistical mechanics. The quasi-harmonic approximation with a cut-off frequency of 100 cm⁻¹ was used to correct inaccuracies of the harmonic approximation for low-frequency vibrational modes.⁵⁶ The third term is a standard-state correction, accounting for the free energy change from 1 mol ideal gas at 1 atm pressure to a concentration in solution of 1 M for all non-MOF species except for the water molecule, where the concentration was changed to that of pure water, 55.34 M,⁵⁷ yielding correction terms of 1.89 and 4.27 kcal/mol, respectively. MOF species were not subject to this correction since they were treated as solids. An additional standard state correction was applied to the hydronium ions, the deprotonated form of the GlyGly dipeptide, and the hydroxyl ions to correct for the pH effect at physiological pH.⁵⁸ To account for the free energy change required to change hydronium and hydroxyl ions from their standard state

conditions, pH 0 and pH 14, respectively, to another pH, the following correction terms were applied:

$$\Delta G'(H^+) = \Delta G^0(H^+) + 2.303RT(\text{pH}^0 - \text{pH}') \quad (2)$$

$$\Delta G'(OH^-) = \Delta G^0(OH^-) + 2.303RT(\text{pH}' - \text{pH}^0) \quad (3)$$

In case of the GlyGly dipeptide, the zwitterionic form is dominant at pH 7. However, since in some cases the deprotonated form was used as a reactant at this pH, the calculated free energy of the deprotonated form was computed at this pH by using equation (4). The alternative approach, computing the free energy of this species from a thermodynamic cycle based on deprotonation by hydroxide ion, was found to be much less accurate.

$$\Delta G^{\text{depr}}(\text{pH}) = 2.303 RT(\text{pK}_a - \text{pH})^{17} \quad (4)$$

Results and Discussions

Nature of the dangling ligands for MOF-808 in water solvent

In the first step, the dangling ligand decoration of the MOF-808 under the conditions used for catalysis in the previous experimental study,²⁶ which we refer to as ‘physiological’ conditions, has been investigated. In pursuit of the most thermodynamically favorable ligation modes of MOF-808 under the biologically relevant conditions, various possible structures differing in the nature and arrangement of the dangling ligands have been considered (a full list is provided in SI, Figure S1). These have been divided into five groups based on the type and arrangement of the coordinated dangling ligands. The first group includes structures containing six formate ions coordinated in either bridging or chelating bidentate fashion (Figure 2a and 2d), denoted as the all-formate bridging and all-formate chelating case, respectively. The second group consists of a series of mixed structures, generated by replacing one, two, three, four or five formate ions in the all-formate structure by in each case one water molecule and one hydroxyl group. These substitutions preserve the charge neutrality and coordinative saturation of the MOF (an example of the singly exchanged case is shown in Figure 2b). In case of a multiple exchange, the relative positions of the water molecule and hydroxyl ions are indicated, so as to distinguish between different isomeric forms,

e.g. mix-(1,2)-water/hydroxyl stands for a mixed structure obtained by replacing two adjacent formate ions for two water molecule/hydroxyl ion pairs. The third group, which we refer to as the mix-2waters case, contains just one structure, an isomer of the singly exchanged mix-water/hydroxyl case. It is obtained by moving one proton (the one closest to the hydroxide) from the inorganic core to the dangling hydroxyl group, creating a $Zr_6O_5(OH)_3$ core and a dangling water molecule instead (Figure 2c). The fourth group also contains just one structure, and is the all-water/hydroxyl case, obtained by replacing each formate group of the all-formate case by one water molecule and one hydroxyl group, arranged to maximize hydrogen bonding (Figure 2f). The fifth group comprises a structure obtained by replacing one formate ion by one water molecule and one hydroxide as ligands, but these are placed *non-adjacently*, separated by a formate ion coordinated in a chelating fashion, as depicted in the Figure 2e, and designated as mix-bridge/chel water/hydroxyl. Considering the known challenges to accurately model solvation of charged species,⁵⁹ relative free energies for ligand exchange reactions were calculated following two reaction models: one reaction model with non-

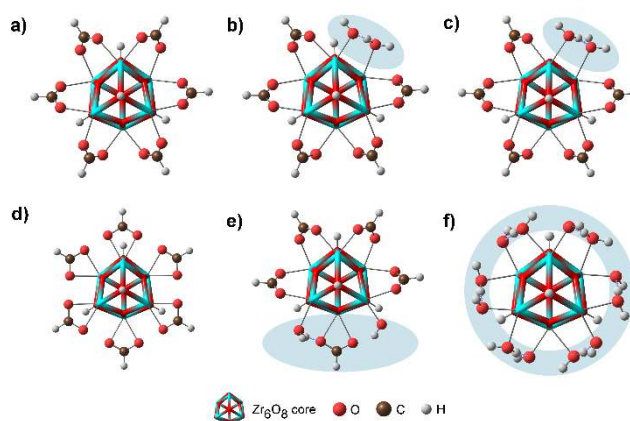
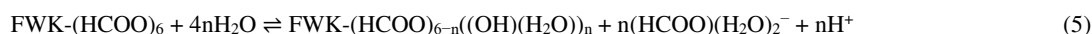


Figure 2 A subset of possible dangling ligand decorations of the single Zr_6 -core unit – modifications made to reach mixed structures are highlighted (benzoate ligands are omitted for clarity). a) all-formate bridging. b) mix-water/hydroxyl. c) mix-2waters. d) all-formate chelating. e) mix-bridge/chel water/hydroxyl. f) all-water/hydroxyl.

conserved (non_con) numbers of charged particles in the reaction process (equation 1) and one reaction model with conserved (con) number of charged particles in the reaction process (equation 2).



where FWK represents the $Zr_6(\mu_3\text{-O})_4(\mu_3\text{-OH})_4(\text{BzO})_6$ framework subunit of the cluster model.

The most relevant relative free energies for different ligand exchange reactions and different isomeric rearrangements are presented in Table 1 (a full list is provided as Figure S1 in the SI). Comparing the two different isomeric forms of the all-formate case, it can be observed that the computed free energy of the chelating isomer is significantly higher than that of the bridging

form. Therefore, the chelating form seems very unlikely to exist in water solvent and can be safely ruled out from future consideration. Contrarily, ligand exchange reaction of formate ions for water molecule/hydroxyl ion pairs appears to be feasible in some cases. Analyzing data for both con and non_con reaction models, two things can be noted. First, replacement of

a formate ion by a water molecule/hydroxyl ion pair is typically an endergonic process under the given reaction conditions. Second, the calculated relative free energies of the corresponding species differ, depending on which reaction model was used to compute them, by a multiple of about 12.5 kcal/mol, the multiplier being the number of formate ions exchanged. This difference can be described as originating from an inaccurate description of water autoprotolysis (note that subtracting reaction (5) from reaction (6) yields the autoprotolysis reaction). This low accuracy is most likely due to limitations of the model used, in particular the implicit solvent model, since the description of solvation of the small charged species involved is known to lead to errors in continuum models.⁵⁹ This problem was expected, and in an attempt to mitigate it, microsolvation of the charged species was used, as indicated in reactions (5) and (6). However, at least with the number of microsolvating water molecules used here, the

problem was not completely eliminated, leading to the noted difference in free energies computed with 'con' and 'non_con' approaches. While improved results might be expected with more microsolvation, adding more water molecules quickly leads to a conformational explosion and difficulties in locating the global minimum for each case, restricting the number of explicit solvent molecules that can practically be used. The present approach represents a reasonable compromise between feasibility and accuracy. We note that reaction models (such as the 'con' model of reaction (6)) with the same number of charged species on the reactants' and the products' side often feature error cancellation, and should therefore give more reliable results.⁵⁹ For this reason, the energetics derived from the 'con' approach, reaction (6), was used in the remainder of this paper because of its favorable error cancellation and reasonable agreement with the experimental findings (see SI for more details).

Table 1. Free energies relative to the free energy of the all-formate bridging form (kcal/mol)

Group	Species composition	Designation	$\Delta G_{\text{con}}^{0***}$	$\Delta G_{\text{non_con}}^{0***}$
I	$\text{Zr}_6(\mu_3\text{-O})_4(\mu_3\text{-OH})_4(\text{BzO})_6(\text{HCOO})_{6b}^*$	all-formate bridging	0	0
	$\text{Zr}_6(\mu_3\text{-O})_4(\mu_3\text{-OH})_4(\text{BzO})_6(\text{HCOO})_{6c}^{**}$	all-formate chelating	38.6	38.6
	$\text{Zr}_6(\mu_3\text{-O})_4(\mu_3\text{-OH})_4(\text{BzO})_6(\text{HCOO})_{5b}(\text{OH})(\text{H}_2\text{O})$	mix-water/hydroxyl	-0.1	12.5
II	$\text{Zr}_6(\mu_3\text{-O})_4(\mu_3\text{-OH})_4(\text{BzO})_6(\text{HCOO})_{4b}(\text{OH})_2(\text{H}_2\text{O})_2$	mix-(1,3)- water/hydroxyl	1.2	26.3
	$\text{Zr}_6(\mu_3\text{-O})_4(\mu_3\text{-OH})_4(\text{BzO})_6(\text{HCOO})_{3b}(\text{OH})_3(\text{H}_2\text{O})_3$	mix-(1,3,5)- water/hydroxyl	3.0	40.7
	$\text{Zr}_6(\mu_3\text{-O})_4(\mu_3\text{-OH})_4(\text{BzO})_6(\text{HCOO})_{2b}(\text{OH})_4(\text{H}_2\text{O})_4$	mix-(1,2,3,4)- water/hydroxyl	2.6	52.8
III	$\text{Zr}_6(\mu_3\text{-O})_4(\mu_3\text{-OH})_4(\text{BzO})_6(\text{HCOO})_{1b}(\text{OH})_5(\text{H}_2\text{O})_5$	mix-(1,2,3,4,5)- water/hydroxyl	3.3	66.1
	$\text{Zr}_6(\mu_3\text{-O})_4(\mu_3\text{-OH})_4(\text{BzO})_6(\text{OH})_6(\text{H}_2\text{O})_6$	all-water/hydroxyl	5.3	80.6
	$\text{Zr}_6(\mu_3\text{-O})_5(\mu_3\text{-OH})_5(\text{BzO})_6(\text{HCOO})_{5b}(\text{H}_2\text{O})_2$	mix-2waters	11.7	24.3
V	$\text{Zr}_6(\mu_3\text{-O})_4(\mu_3\text{-OH})_4(\text{BzO})_6(\text{HCOO})_{4b}(\text{OH})(\text{HCOO})_c(\text{H}_2\text{O})$	mix-bridge/chel water/hydroxyl	8.4	17.8

* Subscript "b" designates bridging binding mode
 ** Subscript "c" designates chelating binding mode
 *** See equations 5 and 6

Results obtained using the 'con' or conserved number of charged particles model from Table 1 indicate that the formate to (water,hydroxide) ligand exchange reactions at physiological pH are roughly thermoneutral in Gibbs energy terms, and can probably not be distinguished from zero within the expected computational errors of our protocol. The first such ligand exchange reaction from the all-formate bridging case has a predicted ΔG_{con}^0 of -0.1 kcal/mol. In order to understand the expected behavior of MOF-808 in neutral water, it is necessary to take into account that starting from the pristine all-formate bridging form, each replacement of a formate ion by 2 water molecules under the given conditions results in a water molecule/hydroxyl ion pair coordination to the MOF, preserving charge neutrality of the MOF. In our calculations, the replacement is accompanied by release of a formate ion and a proton to the liquid phase. This will induce a drop in pH, which in turn will impede further ligand exchange. Based on these calculations, dissolving MOF-808 in pure water is predicted to yield a weakly acidified solution and a MOF with partial exchange of the formate ligands and a mixed structure. This is indeed in line with experiment: mixing MOF-808 with pure water solvent has been observed to lead to a drop in pH to

3.84 (see SI for details of this experiment). It is also in line with a previous study showing increased acidity of water molecules bonded to zirconium hydroxide clusters.⁶⁰ Following the insights from the experimental observations regarding the pH drop, calculations of the energetics of the ligand exchange reactions at the newly established pH value of 3.84 were done. Here, a significant increase in relative free energy of about 4 kcal/mol per exchanged formate was obtained (see Table S6 in SI for details). Due to this high increase in relative free energy for ligand exchange reactions at the lower pH value, multiple ligand exchange becomes energetically very demanding, implying that only single exchange may take place at this pH value. This is in line with previous experimental observations on the structure of the MOF-808, where a singly exchanged species was observed.^{30,31} However, in contrast to these reports, where the released proton was proposed to originate from the inorganic core, leaving 2 coordinated water molecules coordinated, our results strongly suggest that proton loss from one of the coordinated water molecules is by about 12 kcal/mol more favorable, leaving the inorganic core intact and a water molecule/hydroxyl ion pair coordinated instead.

Binding modes of GlyGly to MOF-808

One of the key steps in the catalyzed hydrolysis of a peptide bond is binding of a peptide to a catalyst, and that is our next point of focus. To derive a better understanding of the energetics of this process, a range of possible binding modes of the GlyGly dipeptide to the MOF-808 core have been explored computationally, considering as a reference point the bridging

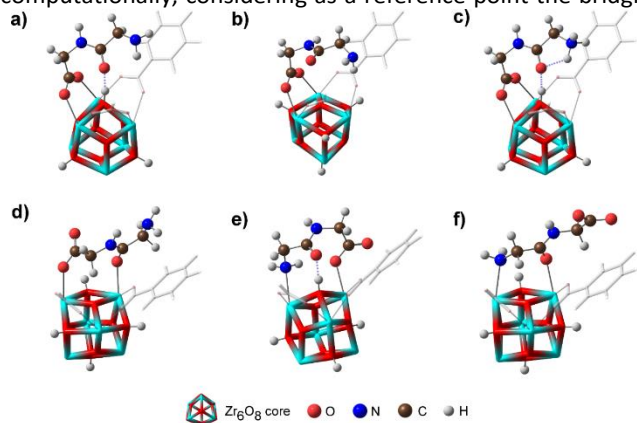


Figure 3 A subset of possible GlyGly – MOF-808 binding modes (only selected ligands of the cluster model are depicted for clarity). a) a1 b) a2 c) a1⁺ d) b1 e) c1 f) d1

all-formate form of the starting MOF-808 structure. Given the insight from the previous section about the possible presence of various mixed and unmixed structures of the MOF-808 at physiological pH, the binding study has been additionally expanded for the all-water/hydroxyl case as a possible starting structure of the MOF-808, to investigate the effect of different possible ligand decorations on the binding energetics.

Coordination of the dipeptide to the MOF is anticipated to occur in a bidentate fashion by replacing a formate ion or a water molecule/hydroxyl ion pair, since bidentate binding was previously shown to be typically substantially more favorable than monodentate binding.⁶¹ At physiological pH, the leaving formate group is assumed to maintain its ionic form in the liquid phase, whereas the leaving molecule/hydroxyl ion pair is expected to abstract a proton and form two water molecules. This proton was modeled to originate from the terminal ammonium group of the zwitterionic dipeptide, leading to formation of an overall neutral MOF-dipeptide complex (Figure 3a,b and 3d-f). In case of substitution to yield a leaving formate ion, two approaches were used. In the first of these, the zwitterionic form of the dipeptide was modelled as replacing a formate ligand, resulting in a positively charged MOF-dipeptide binding complex (Figure 3c). In the second approach, the deprotonated form of the zwitterionic dipeptide was modelled as replacing a formate ligand, resulting in a neutral MOF-dipeptide binding complex. In this approach, the required computed free energy of the anionic form of the dipeptide at physiological pH was obtained as previously described in the Computational details section.

Apart from the overall charge of the complex, and the nature of the ligand undergoing substitution, there are further aspects to take into account when modelling dipeptide binding. The most important of these is that chelating dipeptide binding can occur through atoms of three different functional groups of the

dipeptide: the carboxylic oxygen atoms, the peptide bond amide oxygen atom, and/or the amine nitrogen atom, with two atoms from these groups taking part in binding to the inorganic core. We have considered the following atom pairs: the two carboxylic oxygen atoms, one of the carboxylic oxygen atoms and the amide oxygen atom, the amide oxygen atom and the amine nitrogen atom, and one of the carboxylic oxygen atoms and the amine nitrogen atom (Figure 3). Complexes formed this way are labeled as a, b, c and d, respectively. A second aspect to take into account is that for each of the binding complexes, there are two possible orientations of the dipeptide with respect to the alternating μ_3 -O/ μ_3 -OH environment of the MOF core, which we refer to here as conformation 1 or 2 (e.g. a1 or a2 from Table 2, corresponding to Figure 3a and 3b structures, respectively). Positively charged complexes are additionally qualified with the corresponding “+” sign in the superscript. Calculated binding free energies of the dipeptide with respect to the bridging all-formate form of the MOF-808 as a reference point are presented in Table 2.

Table 2. Relative free energies of MOF-808-GlyGly complex formation (kcal/mol)

Species composition	Label	ΔG^0
$Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-COO-GlyGly_1$	a1	-5.7
$Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-COO-GlyGly_2$	a2	-2.5
$Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-COO-GlyGly_1^+$	a1 ⁺	1.3
$Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-COO-GlyGly_2^+$	a2 ⁺	1.2
$Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-CO/COO-GlyGly_1$	b1	-4.9
$Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-CO/COO-GlyGly_2$	b2	-0.5
$Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-CO/COO-GlyGly_1^+$	b1 ⁺	11.9
$Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-CO/COO-GlyGly_2^+$	b2 ⁺	24.2
$Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-COO/NH_2-GlyGly_1$	c1	-2.5
$Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-COO/NH_2-GlyGly_2$	c2	1.3
$Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-CO/NH_2-GlyGly_1$	d1	1.7
$Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-CO/NH_2-GlyGly_2$	d2	5.6

Three different factors that determine binding affinity may be inferred from Table 2. First, formation of neutral complexes is more favorable than formation of their positively charged counterparts, indicating a preference of the dipeptide to release a proton associated with the binding process. This is not surprising since the dipeptide replaces a negatively charged group during the binding process. Second, the group that forms the strongest interaction with the Zr_6 core is the carboxylic acid group, followed by the carbonyl group, then the amine group. Accordingly, bidentate binding of the carboxylic group (similar to that of the departing formate ligand) is most strongly favoured, while binding through the carbonyl oxygen and the amine nitrogen is least favourable lying a predicted 7.4 kcal/mol higher in free energy. Third, the relative orientation of the dipeptide and of the alternating pattern of μ_3 -O and μ_3 -OH groups within the Zr_6 core plays a non-negligible role in complex stability. For example, in the case of dipeptide complexes involving both oxygen atoms of the carboxylic group, a conformation in which the position relative to the μ_3 -O and μ_3 -OH groups allows formation of a hydrogen bond between one μ_3 -OH group and the amide oxygen atom is more stable by

about 3 kcal/mol compared to one where this is not possible. In other cases, the dipeptide is accommodated closer to the inorganic core, featuring possible steric effects between μ_3 -OH hydrogen atom and a hydrogen atom bonded to an α -carbon atom, resulting in different binding affinities of about 4 kcal/mol for different conformations.

Subsequent consideration of dipeptide binding to the all-water/hydroxyl forms as a possible starting structure of the MOF-808 was narrowed to those cases where only neutral binding complexes are formed in their most stable conformations. As mentioned earlier in this subsection, these reactions include an incoming zwitterionic form of the dipeptide, with release of two neutral water molecules and with formation of neutral forms of the MOF species, so these reactions only include neutral species. Binding free energies for the cases of the all-water/hydroxyl MOF-808 starting structures, together with the all-formate case, are summarized in Table 3.

Table 3. Relative free energies of MOF-808-GlyGly complex formation for different ligand decorations (kcal/mol)

Species composition	L = HCOO ⁻	L = (OH)(H ₂ O) ⁻
Zr ₆ (μ_3 -O) ₄ (μ_3 -OH) ₄ (BzO) ₆ L ₅ -COO-GlyGly_1	-5.7	-0.7
Zr ₆ (μ_3 -O) ₄ (μ_3 -OH) ₄ (BzO) ₆ L ₅ -CO/COO-GlyGly_1	-4.9	4.1
Zr ₆ (μ_3 -O) ₄ (μ_3 -OH) ₄ (BzO) ₆ L ₅ -COO/NH ₂ -GlyGly_1	-2.5	10.4
Zr ₆ (μ_3 -O) ₄ (μ_3 -OH) ₄ (BzO) ₆ L ₅ -CO/NH ₂ -GlyGly_1	1.7	10.5

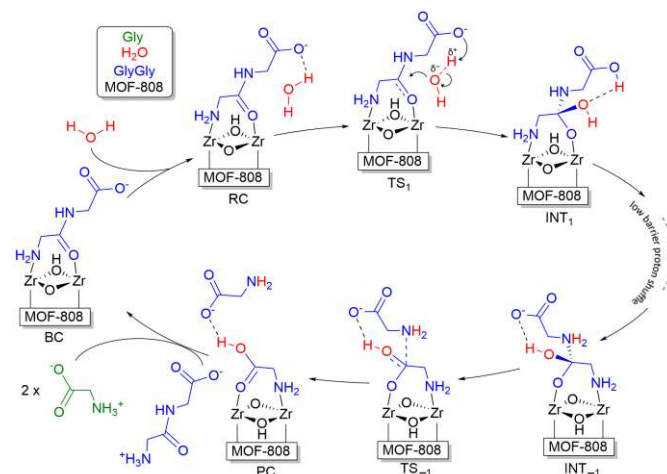
Regardless of the MOF-808 ligand decoration, the same trend in binding affinities with respect to the anchoring groups involved in binding is observed, as evident from Table 3. However, as was the case with the all-formate bridging and all-water/hydroxyl structures, binding complexes decorated with 5 formate capping ligands were characterized by a higher stability than their counterparts with 5 water molecules and hydroxyl ions. Therefore, the most promising candidates for the subsequent mechanistic study were selected from the Zr₆(μ_3 -O)₄(μ_3 -OH)₄(BzO)₆(HCOO)₅-GlyGly group of binding complexes.

Reaction mechanism of the MOF-808 catalyzed hydrolysis of the GlyGly dipeptide bond

In this third part of our study, some plausible reaction pathways for the MOF-808 catalyzed hydrolysis have been investigated in detail. As starting points the binding complexes that involve amide oxygen coordination to the inorganic node have been selected as promising candidates, since they are expected to lead to activation of the inert peptide bond, making it more susceptible to hydrolysis. This means that we assume that the most stable binding form of the dipeptide substrate, a complex involving bidentate anchoring of the carboxylic group to the inorganic node, is not likely to lead directly to hydrolysis since it leaves the dipeptide bond essentially unactivated. Hence only those binding complex structures with the dipeptide coordinated to the MOF through CO/NH₂ or CO/COO groups

have been chosen for further mechanistic study, because of their feasible binding and initial activation of the dipeptide bond through coordination of the amide oxygen atom to a zirconium atom.

In the first step of hydrolysis, a nucleophilic attack on the amide carbon is anticipated. The nucleophile may in principle originate from the core oxygen atoms of the Zr₆O₈ clusters, from a coordinated water molecule/hydroxyl ligand groups of mixed structures, or from the solvent water molecules. Inspection of the complex structures in Figure 3 shows that the core oxygen atoms are not close enough to the carbonyl carbon to act as nucleophiles, so even leaving aside their likely low nucleophilicity, their involvement has been assumed to be impossible. Likewise, nucleophilic attack by dangling coordinated hydroxide ions or water molecules appears to be very unfavorable due to prohibitive distances to the amide carbon. Hence only attack by an external water molecule has been considered in detail. Also, for this reason, due to the considerable computational expense, we decided that the Zr₆(μ_3 -O)₄(μ_3 -OH)₄(BzO)₆(HCOO)₅-GlyGly complex types would not be considered in the mechanistic study, since the mechanisms with external water molecule nucleophiles can be equally well described based on use of Zr₆(μ_3 -O)₄(μ_3 -OH)₄(BzO)₆(HCOO)₅-GlyGly models. These are also lower in



Scheme 1 Modelled catalytic cycle for MOF-808-induced hydrolysis of the GlyGly dipeptide with terminal COO⁻ as a general base

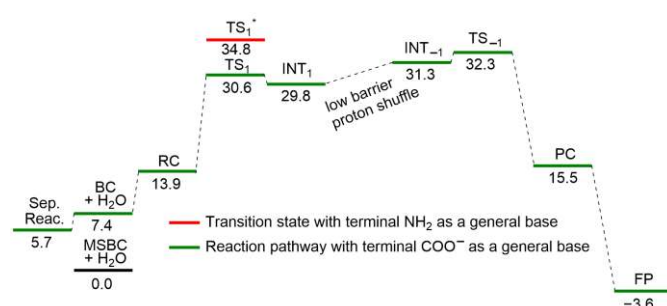


Figure 4 Calculated free energies (kcal/mol) along the hydrolysis pathway relative to the most stable binding complex (MSBC)

relative free energy.

The reaction mechanisms considered in this paper are shown in Scheme 1 and Figure 4. Scheme 1 shows the preferred mechanism, in which an oxygen atom from the carboxylic acid group plays a role as a general base. Figure 4 shows the relative free energy of the key species for the preferred mechanism and for an alternative mechanism in which instead the amino nitrogen atom acts as general base. The starting point in the mechanism is separate reactants, followed by the complex formed upon binding of the dipeptide to the Zr_6 cluster, with coordination of the CO and NH_2 groups, here labelled as 'Sep. Reac.' and 'BC.', respectively. Then a pre-reactive complex is formed between the Zr_6 /dipeptide reactant and a solvent water molecule (RC in Scheme 1). Despite formation of a hydrogen bond, this process is computed to be unfavorable in free energy terms, by 13.9 kcal/mol. This is in part due to the fact that RC involves a less stable binding mode of the dipeptide. Note that given that all reported energies include a continuum SMD model of water solvent, non-specific interactions between water molecules and the Zr_6 /dipeptide reactant are also present prior to forming RC.

The reaction then proceeds in a step involving concerted water molecule dissociation and nucleophilic attack of the forming hydroxyl ion on the amide carbon, passing through TS_1 , the first transition state. Water molecule dissociation is facilitated by proton abstraction by a general base, which is the terminal carboxylic group in the case of TS_1 in Scheme 1, where nucleophilic attack occurs to $Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-CO/NH_2-GlyGly$. However, the general base can in principle also be the terminal amine group in the case of addition to $Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-CO/COO-GlyGly$. As shown in Fig. 4, however, the former type of TS is calculated to be 4.2 kcal/mol lower in free energy, lying just under 31 kcal/mol above the most stable binding complex $Zr_6(\mu_3-O)_4(\mu_3-OH)_4(BzO)_6(HCOO)_5-COO-GlyGly_1$ (MSBC). Given the significant difference in the predicted free energies for these two isomeric TSs, we consider that it is most likely that hydrolysis occurs through the first type of mechanism, with CO/ NH_2 substrate binding and deprotonation through the dangling carboxylate group acting as general base. This is similar to what was found in a previous study on a zirconium-substituted polyoxometalate catalyzed reaction, where the TS corresponding to the nucleophilic attack on the amide carbon was found to be about 20 kcal/mol more stable if general base activation of the nucleophile was performed by the terminal carboxylate group instead of the terminal amine group.¹⁷

After TS_1 , the mechanism is suggested to form the tetrahedral intermediate INT_1 , lying only very slightly lower in free energy. This intermediate is stabilized by direct coordination of the anionic oxygen derived from the carbonyl group to a zirconium center. As in previous computational studies of peptide hydrolysis mechanisms, we expect that after formation of INT_1 , a number of steps will occur involving proton shuffling and conformational changes within the tetrahedral intermediate binding mode. These steps were previously shown to exhibit barriers lower than the rate-limiting step.^{17,62} Accordingly, and given the significant computational challenges associated with modelling reactivity in the present system, these steps were

simply assumed to display similar behavior in the present case and were omitted from the detailed computational study. The suggested end-point of these steps is labelled INT_{-1} in Scheme 1 and Fig. 4, as it is the intermediate immediately preceding product release. In this intermediate, the proton derived from deprotonation of the nucleophilic water molecule has been ultimately moved to the amide nitrogen, whose basicity is greatly enhanced in the tetrahedral intermediate. INT_{-1} is predicted to lie very close in relative free energy to INT_1 , well within the uncertainty of the computational approach.

From INT_{-1} , the last major step in the mechanism, involving C-N bond breaking and re-formation of a carbonyl group, can occur. As often found in peptide hydrolysis mechanisms, the corresponding TS, TS_{-1} , is found to lie close in free energy to the TS_1 for the initial nucleophilic attack. Once again, the terminal carboxylate is found to be able to play the role of general base in TS_{-1} , allowing proton transfer from the forming Zr-coordinated carboxylic acid group. In this case, alternative TSs in which the proton is instead transferred to the amino group have not been considered. The final products of the hydrolysis reaction were theoretically found to have a relative free energy of -9.3 kcal/mol as compared to the starting reactants. Considering the experimental value for the same reaction of -3.6 kcal/mol,⁶³ the calculated value is in fair although not excellent agreement with experiment, bearing in mind the estimated errors of the method, especially concerning solvation.

Likewise, for the two intermediate quantities for which comparison to experiment is possible, the agreement reached here is fair. The first such quantity is the measured value for the Michaelis constant (K_m) of 5.42 mM. At the experimental conditions of 60 °C, this value of K_m corresponds to a binding free energy of the dipeptide to the MOF of -3.4 kcal/mol.²⁶ Comparing this value to the calculated relative free energy of binding for the most stable binding complex of -5.7 kcal/mol, the agreement is acceptable. The second quantity is the rate constant, with k_{max} reported as $11 \times 10^{-4} s^{-1}$ which corresponds to an activation free energy of 24 kcal/mol at 60 °C. This value assumes that all of the Zr_6 cores in the MOF are catalytically active. The present calculations (Fig. 4) suggest an activation free energy of 30.6 kcal/mol, in error by just over 6 kcal/mol. This sort of error, like that for the overall reaction energy mentioned in the previous paragraph, and like that for the water autoprotolysis reaction discussed in a previous section, is obviously quite large, indicating that the protocol used here to describe solvation free energies is still far from being quantitatively accurate. Still, by comparison to similar results obtained for polar reactions in protic solvents,⁶⁴ the agreement with experiment is within the practically expected error bars of the theoretical protocol used. We note that the complex MOF and water environment has been treated here using a continuum SMD model for water, and this is an approximation. While the calculations cannot prove that the suggested mechanism is correct, it certainly seems consistent with the observed properties of the system.

Conclusions

In this study, we set out to gain a better understanding of MOF-808 catalyzed glycine-glycine dipeptide bond hydrolysis by means of DFT calculations. However, gaining mechanistic insight requires being able to characterize the speciation of the Zr₆ nodes under the aqueous reaction conditions, in terms of the ligands that bind to the Zr₆O₈H_{4-x} core – a topic for which there is conflicting experimental evidence. Thus, different MOF-808 ligand decoration types were tested for stability under the reaction conditions. Our results first of all strongly suggest that the core has a strong preference for a Zr₆O₈H₄ protonation state, surrounded by enough negatively-charged ligands to reach overall charge neutrality. At neutral pH, we furthermore conclude that the most stable forms of the MOF are predominantly decorated by bridging formate ligands resulting from the MOF synthesis. However, some exchange of these formate ligands for water and hydroxide ions is predicted to occur. Overall, the Zr₆(μ₃-O)₄(μ₃-OH)₄(BTC)₂(HCOO)₆ and Zr₆(μ₃-O)₄(μ₃-OH)₄(BTC)₂(HCOO)₅(OH)(H₂O) structures are predicted here as the most favorable ones in aqueous solution under physiological pH conditions.

The peptide hydrolysis is predicted to occur through a mechanism in which the dipeptide first substitutes one of the formate ligands to bind to the Zr₆ core. The preferred binding mode involves bridging binding of the dipeptide carboxylate group, as in the substituted formate ligands (and the framework BTC carboxylates). However, isomeric binding modes in which the amide carbonyl oxygen is instead interacting with the metal core are not much higher in free energy and are clearly more likely to undergo further reaction, in a Curtin-Hammett framework. The bound dipeptide is then predicted to participate in a traditional nucleophilic substitution reaction at the carbonyl group through a tetrahedral intermediate which is stabilized by interactions between the amide carbonyl oxygen and a zirconium center. Despite their possible presence, water and hydroxyl ligands on the Zr₆ cluster were not found to facilitate binding of the dipeptide to the MOF, and analysis of the structure of complexes with such water or hydroxide ligands indicated that they could not act as nucleophiles in the hydrolysis step. The reaction instead involves attack by an external water molecule. This water molecule needs to be deprotonated during nucleophilic attack, and our calculations predict that the dipeptide carboxylate group acts as a general base to effect this deprotonation and thereby substantially lowers the barrier for the key addition step. An alternative TS structure in which the dipeptide amine group instead plays the role of general base is higher in free energy. In conclusion, the insight provided in this work on the structure of Zr₆ core at physiological pH, combined with the detailed mechanism of peptide bond hydrolysis catalyzed by a stable metal-organic framework, further contribute to the development of Zr₆-based MOFs as a novel class of artificial proteases.

Conflicts of interest

“There are no conflicts to declare”.

Acknowledgements

The authors thank Hong Giang Ly Thi for performing the experiments relating to the pH drop upon mixing MOF-808 and water. DC thanks L. Gagliardi, D. Truhlar, J. Hupp and Gagliardi group for informative discussions about MOFs. The authors thank KU Leuven and FWO (Research Foundation Flanders) for support. Computational resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation - Flanders (FWO) and the Flemish Government.

Notes and references

- 1 R. Aebersold, M. Mann, *Nature*, **2003**, *422*, 198–207.
- 2 C. Delahunty, J. R. Yates III, *Methods*, **2005**, *35*, 248–255.
- 3 B. Meyer, D. G. Papisotiriou, M. Karas, *Amino Acids*, **2011**, *41*, 291–310.
- 4 D. L. Swaney, C. D. Wenger, J. J. Coon, *J. Proteome Res.*, **2010**, *9*, 1323–1329.
- 5 T. Bolumar, U. Bindrich, S. Toepfl, F. Toldrá, V. Heinz, *Meat Sci.*, **2014**, *98*, 759–765.
- 6 M. L. Katz, J. R. Coates, C. M. Sibigroth, J. D. Taylor, M. Carpentier, W. M. Young, F. A. Wininger, D. Kennedy, B. R. Vuilleminot, C. A. O'Neill, *J. Neurosci. Res.*, **2014**, *92*, 1591–1598.
- 7 J. Polaina, A. P. Maccabe, *Peptidases. Industrial Enzymes Structure, Function and Applications*, Springer, Dordrecht, The Netherlands, 2007.
- 8 S. K. Rai, A. K. Mukherjee, *Biochem. Eng. J.*, **2011**, *54*, 47–56.
- 9 A. Radzicka, R. Wolfenden, *J. Am. Chem. Soc.*, **1996**, *118*, 6105–6109.
- 10 L. R. Croft, *Handbook of Protein Sequence Analysis: A Compilation of Amino Acid Sequences of Proteins with an Introduction to the Methodology*, 2nd ed, John Wiley & Sons, Chichester, United Kingdom, 1980.
- 11 L. Zhu, R. Bakhtiar, N. M. Kostić, *JBC J. Biol. Inorg. Chem.*, **1998**, *3*, 383–391.
- 12 A. Sap, E. De Zitter, L. Van Meervelt, T. N. Parac-Vogt, *Chem. - A Eur. J.*, **2015**, *21*, 11692–11695.
- 13 H. G. T. Ly, G. Absillis, T. N. Parac-Vogt, *Eur. J. Inorg. Chem.*, **2015**, *2015*, 2206–2215.
- 14 H. G. T. Ly, G. Absillis, T. N. Parac-Vogt, *New J. Chem.*, **2016**, *40*, 976–984.
- 15 T. T. Mihaylov, H. G. T. Ly, K. Pierloot, T. N. Parac-Vogt, *Inorg. Chem.*, **2016**, *55*, 9316–9328.
- 16 V. M. Jayasinghe-Arachchige, Q. Hu, G. Sharma, T. J. Paul, M. Lundberg, D. Quinonero, T. N. Parac-Vogt, R. Prabhakar, *J. Comput. Chem.*, **2018**, *40*, 51–61.
- 17 H. G. T. Ly, T. Mihaylov, G. Absillis, K. Pierloot, T. N. Parac-Vogt, *Inorg. Chem.*, **2015**, *54*, 11477–11492.
- 18 J. R. Long, O. M. Yaghi, *Chem. Soc. Rev.*, **2009**, *38*, 1213–1214.
- 19 H. Furukawa, K. E. Cordova, M. O'keeffe, O. M. Yaghi, *Science*, **2013**, *341*, 1230444/1–12.
- 20 G. Férey, *Chem. Soc. Rev.*, **2008**, *37*, 191–214.
- 21 J. Lee, O. K. Farha, J. Roberts, K. A. Scheidt, S. T. Nguyen, J. T. Hupp, *Chem. Soc. Rev.*, **2009**, *38*, 1450–1459.
- 22 B. Kesanli, W. Lin, *Coord. Chem. Rev.*, **2003**, *246*, 305–326.
- 23 L. Ma, C. Abney, W. Lin, *Chem. Soc. Rev.*, **2009**, *38*, 1248–1256.
- 24 D. Farrusseng, S. Aguado, C. Pinel, *Angew. Chemie Int. Ed.*, **2009**, *48*, 7502–7513.
- 25 H.-C. J. Zhou, S. Kitagawa, *Chem. Soc. Rev.*, **2014**, *43*, 5415–5418.
- 26 H. G. T. Ly, G. Fu, A. Kondinski, B. Bueken, D. De Vos, T. N. Parac-Vogt, *J. Am. Chem. Soc.*, **2018**, *140*, 6325–6335.

- 27 S.-Y. Moon, Y. Liu, J. T. Hupp, O. K. Farha, *Angew. Chemie Int. Ed.*, **2015**, *54*, 6795–6799.
- 28 J. E. Mondloch, M. J. Katz, W. C. Isley III, P. Ghosh, P. Liao, W. Bury, G. W. Wagner, M. G. Hall, J. B. DeCoste, G. W. Peterson, R. Q. Snurr, C. J. Cramer, J. T. Hupp, O. K. Farha, *Nat. Mater.*, **2015**, *14*, 512–516.
- 29 M. R. Momeni, C. J. Cramer, *ACS Appl. Mater. Interfaces*, **2018**, *10*, 18435–18439.
- 30 J. Jiang, F. Gándara, Y.-B. Zhang, K. Na, O. M. Yaghi, W. G. Klemperer, *J. Am. Chem. Soc.*, **2014**, *136*, 12844–12847.
- 31 C. A. Trickett, T. M. Osborn Popp, J. Su, C. Yan, J. Weisberg, A. Huq, P. Urban, J. Jiang, M. J. Kalmutzki, Q. Liu, J. Baek, M. P. Head-Gordon, G. A. Somorjai, J. A. Reimer, O. M. Yaghi, *Nat. Chem.*, **2019**, *11*, 170–176.
- 32 E. Plessers, G. Fu, C. Tan, D. De Vos, M. Roefsaers, *Catalysts*, **2016**, *6*, 104/1-11.
- 33 A. L. Loosen, F. de Azambuja, S. Smolders, J. Moons, C. Simms D. E. ; De Vos, T. Parac-Vogt, *Chem. Sci.*, **2020**, DOI: 10.1039/D0SC02136A.
- 34 Y. Bai, Y. Dou, L.-H. Xie, W. Rutledge, J.-R. Li, H.-C. Zhou, *Chem. Soc. Rev.*, **2016**, *45*, 2327–2367.
- 35 H. Furukawa, F. Gándara, Y.-B. Zhang, J. Jiang, W. L. Queen, M. R. Hudson, O. M. Yaghi, *J. Am. Chem. Soc.*, **2014**, *136*, 4369–4381.
- 36 X. Yan, K. Wang, X. Xu, S. Wang, Q. Ning, W. Xiao, N. Zhang, Z. Chen, C. Chen, *Inorg. Chem.* **2018**, *57*, 8033–8036.
- 37 N. Planas, J. E. Mondloch, S. Tussupbayev, J. Borycz, L. Gagliardi, J. T. Hupp, O. K. Farha, C. J. Cramer, *J. Phys. Chem. Lett.*, **2014**, *5*, 3716–3723.
- 38 A. D. Becke, *Phys. Rev. A*, **1988**, *38*, 3098–3100.
- 39 J. P. Perdew, *Phys. Rev. B*, **1986**, *33*, 8822–8824.
- 40 S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.*, **2011**, *32*, 1456–1465.
- 41 S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.*, **2010**, *132*, 154104/1-19.
- 42 F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.*, **2005**, *7*, 3297–3305.
- 43 D. Andrae, U. Häußermann, M. Dolg, H. Stoll, H. Preuß, *Theor. Chim. Acta*, **1990**, *77*, 123–141.
- 44 S. Ling, B. Slater, *Chem. Sci.*, **2016**, *7*, 4706–4712.
- 45 F. Neese, *J. Comput. Chem.*, **2003**, *24*, 1740–1747.
- 46 K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, (Chem. Phys. Letters 240 (1995) 283-290). *Chem. Phys. Lett.*, **1995**, *242*, 652–660.
- 47 K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs, *Theor. Chem. Acc.*, **1997**, *97*, 119–124.
- 48 F. Weigend, *Phys. Chem. Chem. Phys.*, **2006**, *8*, 1057–1065.
- 49 F. Neese, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2012**, *2*, 73–78.
- 50 F. Neese, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2018**, *8*, e1327. doi: 10.1002/wcms.1327.
- 51 A. D. Becke, *Cit. J. Chem. Phys.*, **1993**, *98*, 5648–5652.
- 52 C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B*, **1988**, *37*, 785–789.
- 53 P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.*, **1994**, *98*, 11623–11627.
- 54 A. K. Dutta, F. Neese, R. Izsák, *J. Chem. Phys.*, **2016**, *144*, 034102/1-12.
- 55 A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B*, **2009**, *113*, 6378–6396.
- 56 R. F. Ribeiro, A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B*, **2011**, *115*, 14556–14562.
- 57 V. S. Bryantsev, M. S. Diallo, W. A. Goddard III, *J. Phys. Chem. B*, **2008**, *112*, 9709–9719.
- 58 R. A. Alberty, *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1994**, *1207*, 1–11.
- 59 J. Ho, M. L. Coote, *Theor. Chem. Acc.*, **2010**, *125*, 3–21.
- 60 M. Åberg, J. Glaser, *Inorg. Chim. Acta*, **1993**, *206*, 53–61.
- 61 M. L. Mendonca, R. Q. Snurr, *ACS Catal.*, **2020**, *10*, 1310–1323.
- 62 Q. Hu, V. M. Jayasinghe-Arachchige, G. Sharma, L. F. Serafim, T. J. Paul, R. Prabhakar, *WIREs Comput. Mol. Sci.*, **2020**, 10:e1466
- 63 R. B. Martin, *Biopolymers*, **1998**, *45*, 351–353.
- 64 Z. Liu, C. Patel, J. N. Harvey, R. B. Sunoj, *Phys. Chem. Chem. Phys.*, **2017**, *19*, 30647–30657.