Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

BIOLOGIA CELLULARE E MOLECOLARE

Ciclo XXIX

Settore Concorsuale di afferenza: 03/B1

Settore Scientifico disciplinare: CHIM/03

BIOCHEMICAL AND STRUCTURAL STUDIES ON UREASE INHIBITION, A NICKEL-DEPENDENT VIRULENCE FACTOR

Presentata da: Luca Mazzei

Coordinatore Dottorato

**Chiar.mo Prof.
Giovanni Capranico**

Relatore

**Chiar.mo Prof.
Stefano Ciurli**

Correlatore

Dott. Francesco Musiani

Esame finale anno 2017

"Today is only one day in all the days that will ever be.
But what will happen in all the other days that ever come
can depend on what you do today."

For Whom the Bell Tolls (1940)

E.H.

TABLE OF CONTENTS

ANNEX 1:   Urease (book chapter)

ANNEX 2:   Fluoride inhibition of *Sporosarcina pasteurii* urease: structure and thermodynamics

ANNEX 3:   Kinetic and structural studies reveal a unique binding mode of sulfite to the nickel center in urease

ANNEX 4:   Inactivation of urease by 1,4-benzoquinone: chemistry at the protein surface

ANNEX 5:   Inactivation of urease by catechol: Kinetics and structure

ACKNOWLEDGMENTS

INTRODUCTION

Enzymes play a major role in most of the biochemical processes and reactions that are at the basis of life. However, enzymatic alterations are involved in many human diseases, making enzymes very attractive by the pharmaceutical industry and academia. Indeed, enzymes are the ideal targets for drug discovery processes aimed to the development of molecules able to modulate their activity and eventually block their negative consequences.

In the present work, the attention has been focused on the nickel-dependent urease as a target enzyme. Urease is found in several organisms, such as bacteria, algae, fungi and higher plants and catalyses the rapid hydrolysis of urea to give ammonia and bicarbonate, causing an overall increase of pH. The relevance of urease derives from the fact that it is a virulence factor for several human pathogenic microorganisms. A paradigm in this sense is *Helicobacter pylori*, a Gram-negative ureolytic bacterium that exploits the urease activity to colonize the human gastric mucosa causing chronic gastritis, peptic and duodenal ulcers. It is the only bacterium that has been classified as a "class 1" carcinogen by the World Health Organization (WHO). Moreover, WHO has recently identified the twelve most important antibiotic-resistant bacteria for which new treatments are urgently required and, among those, nine organisms are ureolytic.[1] Urease is also involved in agro-environmental issues. Being present in the soil in large quantities, it substantially decreases the efficacy of urea-based fertilizers used in agriculture because of ammonia volatilization, thus causing economic and environmental problems. For these reasons, the scientific community has devoted intense efforts in the last several decades for the development of efficient and specific inhibitors of urease able to counteract its negative effects both from a medical and an agro-environmental point of view. The overall conserved tridimensional structure of ureases among the species, together with the

---

[1] (www.who.int/medicines/publications/global-priority-list-antibiotic-resistant-bacteria/en/)

availability of high-resolution X-ray crystal structures of ureases from different organisms, strongly enhances the significant role that structure-based drug design can have in the process of discovery of new urease inhibitors, both as drugs and as molecules to be used as nitrogen stabilizers in agriculture.

In this study, a combination of biochemistry and structural biology has been applied to the urease system to determine the inhibition mode of several known urease inhibitors, with the aim to fully characterize both their interactions and binding efficacy towards the enzyme. All the results shown in here will be useful to develop novel and more efficient urease inhibitors, necessary to modulate its activity and to counterbalance its negative effects.

As the reader will notice, a relevant portion of this dissertation is included as Annexes. In particular, a complete description of urease is given in the published book chapter available as Annex 1. This book chapter should be intended as an introduction section, to be read together with Chapter 1. The reader will be also guided through the other Annexes present in this work during the reading of the thesis, each representing a scientific article published as results of the studies carried out by me, together with other collaborators, during the course of the Doctoral Program in Cellular and Molecular Biology.

CHAPTER 1

STRUCTURE-BASED DRUG DESIGN IN DRUG DISCOVERY

1.1 Drug discovery: past and present

In the past, development of medicines was usually driven by the knowledge acquired on endogenous ligands without a simultaneous awareness on the structure of their cognate biological receptors. Such a paradigm can be highlighted by the emblematic example of steroid hormones. In fact, at the beginning of the 20th century the important role of steroidal molecules in many biological aspects was proved on a biochemical basis. Later, in 1932, their overall tridimensional structure was determined by X-ray crystallography.[1] The resulting combination of the biochemical and structural information available on steroids allowed understanding how the substituted groups present at specific positions around the scaffold provide specific biological activity, endorsing the development of new drugs.[2, 3]

The first idea of modern drug discovery was what we now call *"classical" pharmacology*.[4] It was based on the phenotypic screening of small molecule libraries to be tested in living cells or whole organisms with the aim of identifying compounds with an appropriate therapeutic effect. Nowadays, with the improvements in genomics, molecular biology and related techniques that allow production of large quantities of purified proteins, this process has been overturned. Databases of small molecules are used to perform high throughput screenings in order to evaluate their ability to interact with proteins potentially involved in diseases. Good candidates resulting from the initial

screening are then tested *in vitro* and *in vivo* to detect their real efficacy. This new workflow is usually referred to as *reverse pharmacology.*[4]

In the last 20 years, advances in computational methods have contributed to the rise of the so-called *structure-based drug design* (SBDD) as the basis of drug discovery in reverse pharmacology (Figure 1). Today, SBDD is a fundamental part of most drug discovery programs carried out by companies, simultaneously becoming a major research field for academic laboratories.[5] As a part of the drug discovery process, it is worth noting that structure-based drug design is important in the discovery of drug leads rather than drug products. Many years may be necessary to improve a drug lead in order for it to be used as a drug effective and tolerated by human body. Additionally, time will also be spent in order for the drug to pass through the clinical trials and possibly reach the market.[6]
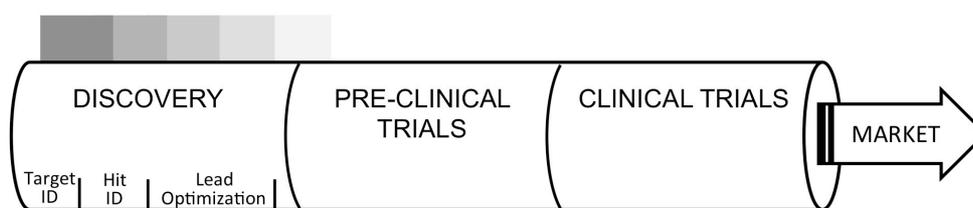
Figure 1: Schematic representation of the drug discovery process. The shaded boxes emphasize where SBDD plays a significant role. Redrawn from: *3D structure and the drug-discovery process.*[6]

Even though SBDD can be theoretically applied to every kind of macromolecule (i.e. proteins, nucleic acids and more complex supramolecular assemblies), in this chapter our discussion will be focused on SBDD performed on enzyme targets. First, a brief introduction on the procedure, as well as a description of the most relevant drug design methods will be given. In the last part of the chapter we will focus on the two peculiar and recurring aspects of the SBDD process, the structural characterization of the target in the absence and in the presence of ligands and the biochemical characterization of the ligand-target interactions.

1.2   An overview of structure-based drug design

Structure-based drug design (SBDD) is an iterative process usually proceeding through multiple steps. The *structure-based* term stands for the need of a tridimensional structure of the target macromolecule, alone and/or in complex with ligands, in order for the drug design process to be tailored at an atomic level. The structural determination of the target can be pursued by using three main methodologies: i) X-ray crystallography, ii) NMR spectroscopy, and iii) homology modelling.[5, 7]

In the first step of a general SBDD approach, the available structural information on the target is exploited to screen small molecule databases by using computer-aided approaches in order to select compounds with potential binding affinity for the target itself. Next, biochemical experiments are carried out in order to corroborate the binding efficacy of the molecules selected by computational procedures (hits) and select those showing an at least micromolar affinity, *in vitro*. In the second step, structure determination of the target in complex with the active hits selected in the previous step is carried out. The latter aspect is necessary to depict, at an atomic level, the mode of interaction between the macromolecule and the testing compounds, on a structural point of view. Moreover, this would show possible portions of the hits that can be improved in order to further optimize the efficacy of ligand binding. In the third step, biochemical and structural information obtained by the previous phases is exploited to chemically modify the hits, and further biochemical experiments are carried out to verify the effects of the chemical modifications in terms of binding efficacy. By iteratively repeating the procedure, an overall optimization of the molecules should be observed in terms of binding properties and specificity for the target.[5, 7]

1.2.1   Identification of a potential binding site on the target

A typical target of a SBDD campaign can be an enzyme involved in a disease. Thus, with the aim of discovering a new drug that would counterbalance its negative effects, the first mandatory step is the determination of the enzyme structure. This can be achieved by using techniques such as X-ray crystallography, NMR spectroscopy and homology modelling,[5, 7] as we will discuss in Section 1.3. Here we can assume that one of these techniques has been successfully applied and the tridimensional structure of the enzyme

under study is already available. In these circumstances, the following step is the exploitation of the structural information for the identification of a potential binding site for the ligand on the target. In the case of enzymes, the most interesting potential binding site for a ligand is certainly the active site. This is in fact the most direct way through which enzyme activity can be suppressed. An alternative strategy can be to identify a region that is not directly involved in the enzyme catalysis but that controls conformational changes that are necessary for the catalysis to occur.[8] Focusing on non-enzyme targets, some interesting arguments can be pointed out. If the subject of our study is a cellular receptor, numerous studies have demonstrated that the binding site for the endogenous ligand is a valuable potential target for the design of cognate ligands, usually small peptides, that bind by exploiting mimicry.[9, 10] Otherwise, if the goal of SBDD is to find a molecule able to modulate a protein-protein interaction, potential binding sites can be the assembly patches that allow this interaction to occur, which are often characterized by relatively compact epitopes consisting of clustered amino acid side chains within the dimensions of a small molecule framework.[11] Moreover, an interesting point is the possibility to use RNA secondary structural elements as target sites. In fact they are usually species specific, able to bind ligands and are often involved in disease states.[12, 13]

1.2.2 Drug design methods

Once a potential binding site on the enzyme under study has been identified by analysing its tridimensional structure, the following objective is the screening of small molecules in order to find a good initial compound that potentially interacts with the target. Small molecules are usually selected by computational methods, such as: i) direct inspection, ii) virtual ligand screening, and iii) *de novo* generation.

Direct inspection relies on a direct structural analysis of ligands that are already known to bind the target macromolecule. If we are working with an enzyme, attention can be focused on substrates, possible cofactors or already established enzyme inhibitors. Direct inspection essentially consists in a chemical modification of such known molecules in order for the interactions between the enzyme binding site and the molecule itself to be maximized, transforming the latter in a potent inhibitor. There are a number of studies demonstrating that analogues of the substrate or modified cofactors can be good candidates for drug development against enzymes.[14-17] Such chemically modified

molecules are preliminary tested *in silico*, by docking procedures that score the interaction between the binding site and the putative ligand.[18-20] Thus, the most promising molecules are synthetized (or purchased, if commercially available) and tested *in vitro* to confirm their binding efficacy.[5, 21]

In virtual ligand screening (VLS), small molecule databases (e.g. *Asinex*, *Chembridge*, *Princeton*, *Enamine* or *Zinc*) are docked to the target binding site, scoring the orientation results according to the predicted interactions.[22] One advantage of this approach is the possibility of refining the database search in order to filter molecules with a specific motif. Molecules that give best results are then purchased and tested *in vitro* to verify their effective properties through biochemical assays. However, databases for VLS do not provide structurally novel molecules, as they are usually already commercially available.[23]

In the *de novo* generation, small fragments or portions of molecules (benzene rings, amino groups, carbonyl groups, etc.) are docked into the targeted binding site. With an iterative process, fragments can be then linked together in order to build a novel molecule that has never been synthesized before and to obtain larger inhibition potency.[22, 24]

These three computational methods can be used alone or can be combined. For example, the experimentally active compounds found after a virtual screening step can be optimized by using a direct inspection approach.

1.3   Structure determination of the target

As introduced at the beginning of this chapter, the prior requirement for any structure-based drug design is the availability of a tridimensional structure of the macromolecule against which we want to develop new drugs. Such a structure will further guide the drug discovery process throughout all the steps of SBDD, playing a significant role in the "ping-pong" procedure that passes through the structural determination and chemical improvements of the testing molecules. We already mentioned the three main methodologies used to provide a structural characterization of the target useful in SBDD, X-ray crystallography, NMR spectroscopy and homology modelling.[5, 7] However, the following discussion will cover only X-ray crystallography and NMR spectroscopy, as they are the two experimental methodologies of choice. They both allow the

determination of high-resolution structures at an atomic level and, when possible, they have to be intended as complementary approaches. Homology modelling consists in the prediction of the tridimensional structure of a protein starting from its amino acid sequence by a structural comparison with one or more homologue proteins of known structure[25, 26] and it is usually used when no tridimensional structure of the protein under study is available. On the other hand, the structures obtained by homology modelling do not have the resolution and the accuracy required for SBDD.

X-ray crystallography has been historically the most common method adopted for the structural data gathering useful in the iterative process of SBDD.[27-29] As a main technique used in our laboratory, a detailed description of X-ray crystallography bases are given in Appendix B. Besides theory, a crystallographic structure suitable as a target for drug design must present some mandatory features.[5] In general, the structural model must be chemically correct, with a final refined structure consistent with the stereochemical parameters. Also, low coordinate errors, as well as B-factors of the atoms belonging to the binding region not greater than the average B-factor for the macromolecule, are needed. A number of tools, usually used prior to a PDB deposition, can be exploited to check the reliability of a structural model, such as PROCHECK,[30] WHATIF[31] or the wwPDB validation web server (http://www.wwpdb.org/).[32] There are some advantages in the use of this technique with respect to NMR spectroscopy. First, X-ray crystallography can be applied to macromolecules virtually without any size limitation.[33] High-resolution crystal structures are available for proteins, nucleic acids, as well as bigger systems such as macromolecular complexes and viruses.[34] Moreover, it enables the direct visualization of the ordered solvent molecules located inside and around the protein model. This information can be used to improve the process of drug discovery. For example, ordered water molecules can be incorporated into the proposed ligand, enhancing ligand-binding properties by increasing the entropy of the system upon water release.[35] Additionally, ordered water molecules can be considered as actual ligands, and interactions between the candidate molecules and water molecules themselves can be maximized.[36] In general, technical improvements regarding X-ray crystallography, such as microfocused beamlines, new generation detectors, and the innovative use of free-electron lasers have reduced the number and dimension of crystals required to determine tridimensional structures. However, there is no denying that the need of a crystal can be a bottleneck in

the characterization of the interaction mode of the ligand with the macromolecule. Some proteins simply do not crystallize. In addition, co-crystallization or soaking can give different results, for example, when proteins require a conformational change for the ligand binding to occur, or when crystal packing negatively affects the ligand binding itself.[37] The non-physiological conditions present in a protein crystal can also give artifacts, even though numerous studies have demonstrated that a macromolecule crystal can be referred to an ordered gel with interstitial spaces through which water and other small molecules freely diffuse.[38]

To complement the information acquirable with X-ray crystallography, as well as to overcome the technique-related issues, NMR spectroscopy is also a valuable technique for drug design purposes. In general, the main advantage of NMR with respect to X-ray crystallography is that the interaction between the target and the ligand can be evaluated in solution and, by using appropriate approaches, this interaction can be characterized on both a structural and kinetic point of view. In particular, NMR experiments can be performed on a protein-ligand mixture by following the resonances of either the ligand or the protein.[39] In the so-called ligand-based NMR methods, no labeling is needed and the detectable resonances, usually pursued through NOE-type experiments[40, 41] and water ligand-observed spectroscopy (Water-LOGSY),[42] belong to the hydrogen atoms of the not bound ligand (the protein-bound ligand behaves like a protein, with a slow tumbling time and undetectable signals). A comparison between the resulting NMR spectra collected for the ligand in the presence or in the absence of the protein can reveal binding information.[43] In protein-based NMR methods, a complete map of the amino acid residues involved in the interaction with the testing compound can be obtained upon a sequence-specific assignment of the backbone resonances. A typical ^{1}H-^{15}N Heteronuclear Single-Quantum Correlation (HSQC)[44] collected for the protein in the absence and in the presence of the molecule would result in a chemical shift perturbation (CSP)[45] of the residues directly involved in the interaction with the ligand. In this context, if the tridimensional structure of the protein is available, the binding site for the ligand on the protein can be easily mapped. Other techniques that do not require a backbone assignment rely on intermolecular NOEs measurements.[46] NMR can overcome X-ray crystallography within the drug design process in some instances, for example when one is studying membrane proteins, which are difficult to crystallize. In this case,

macromolecule-ligand interactions can be evaluated with NMR by mixing in solution natural or artificial membranes and allowing the protein resonances to be detectable.[47] However, the main limitation of NMR spectroscopy is in the size of the macromolecules that can be studied, with sensitivity losses and spectral complexity that allow to work only with macromolecules up to 50-60 kDa.[48] This is an addition to the limited resolution of NMR structures with respect to X-ray structures.

1.4   Biochemical assays on the testing compounds

After the first computational screening of small molecules potentially interacting with the target enzyme, as well as during each step of lead optimization, an *in vitro* biochemical analysis on the testing molecules has to be performed in order to determine a structure-activity relationship (SAR) of the interaction between ligand and target.[49] A typical evaluation flowchart for the determination of the binding mode of a ligand towards an enzyme is schematically presented in Figure 2. Most of the procedures listed in the scheme and described below are examined in detail in Appendix A, hence the following discussion will resume them in order to depict an overview on how they are involved in the SBDD process.
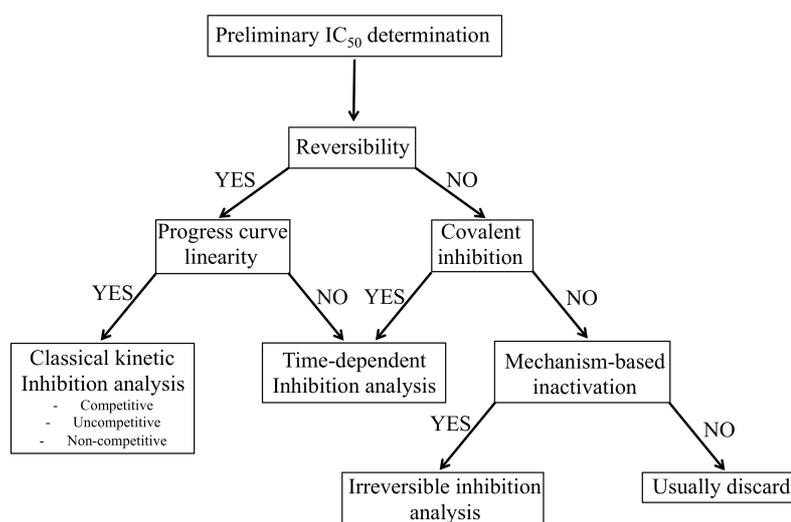
Figure 2: Evaluation flowchart for the compounds tested in the optimization phase of a structure-based drug design campaign. Redrawn from: *Lead optimization and structure-activity relationships for reversible inhibitors.*[49]

In general, a preliminary evaluation of ligand efficacy is made by a concentration-response analysis in order to determine the IC_{50}, namely the concentration of inhibitor that causes a loss of enzyme activity equal to 50%. By measuring the reaction velocity of our target enzyme in the absence (v_0) and in the presence of increasing amounts of inhibitor (v_i) (at fixed concentrations of enzyme and substrate), it would result in a progressive decrease of reaction velocity by increasing the concentration of inhibitor, following a sigmoidal behaviour, until no residual activity would be detectable. A plot of percentage residual activity (v_i/v_0 %) as a function of inhibitor concentration can be drawn as in Figure 3, and IC_{50} can thus be calculated as the mid-point in the semi-logarithmic plot by the following equation:

$$\text{residual activity \%} = \frac{100}{1+\left([I]/IC_{50}\right)} \tag{2.1}$$
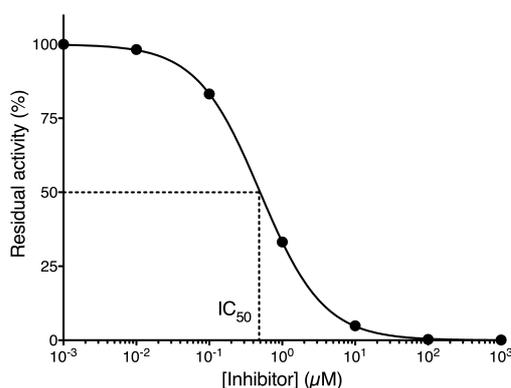
Figure 3: Dose-response semi-log plot of enzyme residual activity as a function of inhibitor concentration. The IC_{50} value for the inhibitor is graphically determined. Redrawn from: *Lead optimization and structure-activity relationships for reversible inhibitors.*[49]

IC_{50} is commonly used to rank-over the relative efficacy of the testing compounds. However, a more exhaustive characterization of the inhibition mechanism can be carried out by determining whether the inhibition is reversible or irreversible and, in the former case, information on the type of reversibility (i.e. rapid or slow) is also achievable. The most common procedure to detect a reversible behaviour is to perform a progress curve experiment by measuring the recovery of enzyme activity after a large dilution of the enzyme-ligand mixture, as described in Figure 4. In such an analysis, a rapid reversible

inhibition would result in almost complete recovery of enzyme activity, with a progress curve profile similar to the one of the enzyme not treated with the inhibitor. A slowly reversible inhibition (slow-binding) would result in a curvilinear progress curve, with a lag phase followed by a linear one, while an irreversible inhibition would otherwise result in a non-recovery of enzyme activity.
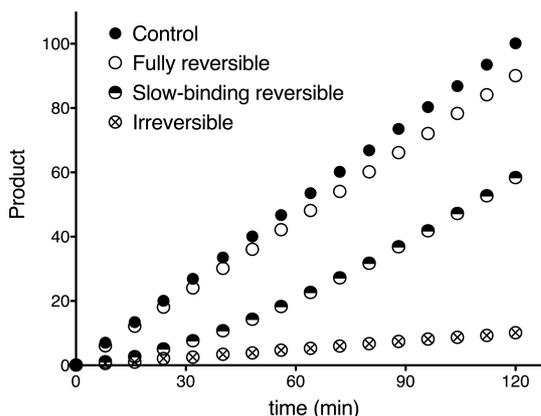
Figure 4: Recovery of enzyme activity after rapid dilution. Analysis of the resulting progress curve provides information on the inhibition mode of the ligand, as described in the main text. Redrawn from: *Lead optimization and structure-activity relationships for reversible inhibitors.*[49]

As a following step, the fine biochemical characterization of the kinetic parameters that regulate the inhibition process has to be performed. In general, a reversible inhibitor is characterized in terms of reversible inhibition type, determining whether it acts with a competitive, uncompetitive or non-competitive reversible inhibition mechanism and determining the equilibrium dissociation constant K_i (M) between the inhibitor and the enzyme. On the contrary, irreversible inhibition can occur through an affinity labeling or a mechanism-based process.[50] Inhibitors that are referred to as affinity labels are molecule with a medium to high intrinsic reactivity towards amino acid residues. In general, such molecules covalently modify the enzyme in a non-specific manner. For this reason a drug design procedure based on this kind of molecules is highly insidious. On the other hand, mechanism-based inhibitors have to be converted in the actual inhibitor species by a catalytic event performed by the enzyme, in order to act either as an affinity label, a transition state analogue, or a tight binding reversible inhibitor.[51] These molecules usually act as competitive inhibitors and are also named suicide substrates. The need of an initial

reaction with the enzyme makes these inhibitors high selective, therefore quite useful as potential drug leads.

Concluding, the fine biochemical characterization of the SAR and the determination of related kinetic parameters is necessary in order to quantitatively measure the binding efficacy of the ligands towards the enzyme, therefore judging by comparison the improvement in lead optimization during each step of the SBDD process.

REFERENCES

1.   O. Rosenheim and H. King (1932), *J. Chem. Technol. Biotechnol.*, **51**:954.

2.   D. J. Ingle (1950), *J. Clin. Endocrinol. Metab.*, **10**:1312.

3.   S. E. Hansen (2008), *Dan. Medicinhist. Arbog*, **36**:109.

4.   T. Takenaka (2001), *BJU Int.*, **88**:7.

5.   A. C. Anderson (2003), *Chem. Biol.*, **10**:787.

6.   R. E. Hubbard (2005), *Mol. BioSyst.*, **1**:391.

7.   C. L. M. J. Verlinde and W. G. J. Hol (1994), *Structure*, **2**:577.

8.   M. Wang, et al. (2003), *J. Biol. Chem.*, **278**:9489.

9.   J. M. Matsoukas, et al. (2000), *Bioorg. Med. Chem.*, **8**:1.

10.   L. Polevaya, et al. (2001), *Bioorg. Med. Chem.*, **9**:1639.

11.   T. R. Gadek and J. B. Nicholas (2003), *Biochem. Pharmacol.*, **65**:1.

12.   A. V. Filikov, et al. (2000), *J. Comput. Aided. Mol. Des.*, **14**:593.

13.   K. E. Lind, Z. Du, K. Fujinaga, B. M. Peterlin and T. L. James (2002), *Chem. Biol.*, **9**:185.

14.   N. A. Roberts, et al. (1990), *Science*, **248**:358.

15.   B. D. Dorsey, et al. (1994), *J. Med. Chem.*, **37**:3443.

16.   M. D. Varney, et al. (1992), *J. Med. Chem.*, **35**:663.

17.   D. C. Chan, C. A. Laughton, S. F. Queener and M. F. Stevens (2001), *J. Med. Chem.*, **44**:2555.

18.   L. Ferreira, R. dos Santos, G. Oliva and A. Andricopulo (2015), *Molecules*, **20**:13384.

19.   J. J. Irwin and B. K. Shoichet (2016), *J. Med. Chem.*, **59**:4103.

20.   S. P. Leelananda and S. Lindert (2016), *Beilstein J. Org. Chem.*, **12**:2694.

21.   V. Lounnas, T. Ritschel, J. Kelder, R. McGuire, R. P. Bywater and N. Foloppe (2013), *Comput. Struct. Biotechnol. J.*, **5**:e201302011.

22.   Q. Li and S. Shah (2017), in *Protein Bioinformatics: from protein modifications and networks to proteomics*, eds. C. H. Wu, C. N. Arighi and K. E. Ross, Springer, New York, NY, pp. 111.

23.   E. Lionta, G. Spyrou, D. K. Vassilatis and Z. Cournia (2014), *Curr. Top. Med. Chem.*, **14**:1923.

24.   W. L. Jorgensen (2004), *Science*, **303**:1813.

25.   J.-M. Chandonia and S. E. Brenner (2005), *Proteins*, **58**:166.

26.   V. K. Vyas, R. D. Ukawala, M. Ghate and C. Chintha (2012), *Indian J. Pharm. Sci.*, **74**:1.

27.   P. Kuhn, K. Wilson, M. G. Patch and R. C. Stevens (2002), *Curr. Opin. Chem. Biol.*, **6**:704.

28.   T. L. Blundell and S. Patel (2004), *Curr. Opin. Pharmacol.*, **4**:490.

29.   L. W. Hardy and A. Malikayil (2003), *Curr. Drug Discov.*, **3**:15.

30.   R. A. Laskowski, M. W. MacArthur, D. Moss and J. M. Thornton (1993), *J. Appl. Crystal.*, **26**:283.

31.   G. Vriend (1990), *J. Mol. Graph.*, **8**:52.

32.   H. Berman, K. Henrick and H. Nakamura (2003), *Nature Struct. Biol.*, **10**:980.

33.   M. Egli (2010), *Curr. Protoc. Nucleic Acid Chem.*, **7**:Unit 7.13.

34.   N. Verdaguer, D. Garriga and I. Fita (2013), in *Structure and physics of viruses: an integrated textbook*, ed. M. G. Mateu, Springer Netherlands, Dordrecht, pp. 117.

35.   P. Y. Lam, et al. (1994), *Science*, **263**:380.

36.   R. A. Powers, F. Morandi and B. K. Shoichet (2002), *Structure*, **10**:1013.

37.   A. Cousido-Siah, et al. (2012), *Proteins*, **80**:2552.

38.   A. McPherson and J. A. Gavira (2014), *Acta Crystallogr. F*, **70**:2.

39.   D. M. Dias and A. Ciulli (2014), *Prog. Biophys. Mol. Biol.*, **116**:101.

40.   M. Mayer and B. Meyer (1999), *Angew. Chem.*, **38**:1784.

41.   M. Mayer and B. Meyer (2001), *J. Am. Chem. Soc.*, **123**:6108.

42.   C. Dalvit, P. Pevarello, M. Tatò, M. Veronesi, A. Vulpetti and M. Sundström (2000), *J. Biomol. NMR*, **18**:65.

43.   P. Śledź, C. Abell and A. Ciulli (2012), in *NMR of Biomolecules*, eds. I. Bertini, K. S. McGreevy and G. Parigi, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, pp. 264.

44.   G. Bodenhausen and D. J. Ruben (1980), *Chem. Phys. Lett.*, **69**:185.

45.   M. P. Williamson (2013), *Prog. Nucl. Magn. Reson. Spectrosc.*, **73**:1.

46.   K. L. Constantine, M. E. Davis, W. J. Metzler, L. Mueller and B. L. Claus (2006), *J. Am. Chem. Soc.*, **128**:7252.

47.   J. Wirmer-Bartoschek and S. Bartoschek (2012), *Future Med. Chem.*, **4**:869.

48.   D. P. Frueh, A. Goodrich, S. Mishra and S. Nichols (2013), *Curr. Opin. Struct. Biol.*, **23**:734.

49.   R. A. Copeland (2013), in *Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists, 2nd Edition*, John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 169.

50.   R. A. Copeland (2013), in *Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists, 2nd Edition*, John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 345.

51.   L. M. Szewczuk, L. Forti, L. A. Stivala and T. M. Penning (2004), *J. Biol. Chem.*, **279**:22727.

CHAPTER 2

THE ACTIVE SITE OF UREASE
AS A BINDING TARGET

2.1 A biochemical and structural characterization of two reversible inhibitors of urease

Two papers have been published by our group in 2014 and 2016, regarding kinetic and structural studies on the inhibition of urease by two inhibitors, fluoride and sulphite.

The model enzyme used for these studies was urease from *Sporosarcina pasteurii* (SPU), a widespread ureolytic soil bacterium. Isothermal titration calorimetry was used in the 6.5–8 pH range to determine the kinetic parameters of the inhibitions, while X-ray crystallography was used to depict, at an atomic level, the structural details of the interactions between the enzyme and the inhibitors.

In the publication describing the inhibition of urease by fluoride, we reported that this ion acts as a pH-dependent inhibitor of urease, with a mixed inhibition mode. In order to provide a structural basis for this kinetic observation, a peculiar structural study was carried out. Indeed, in order to distinguish the iso-electronic fluoride ions and solvent molecules with significantly meaningful statistics, ten X-ray crystal structures of SPU were individually determined, using five crystals of the native enzyme and five crystals of the protein co-crystallized in the presence of fluoride, an unprecedented approach. The analysis of these structures revealed the presence of two fluoride ions coordinated to the Ni(II) metal centres in the active site, a result that is fully consistent with the mixed inhibition mode pointed out by the kinetic experiments. In particular, one fluoride ion replaces the water molecule (W1) bound to Ni(1), thus acting as a competitive inhibitor

by interfering with the binding of the substrate urea to the same metal ion in the first step of the catalysis, while another fluoride ion replaces the hydroxide ion bridging the two Ni(II) ions, interfering with the nucleophilic attack on the substrate urea and thus acting as an uncompetitive inhibitor. The analysis of the hydrogen-bonding network around the nickel-bound fluoride ions further provides a basis for the interpretation of the pH-dependence of inhibition.

In the paper in which the inhibition of urease by sulphite was characterized, we reported a pH-dependent and competitive inhibition mode of sulphite, with an inhibition constant of 0.19 mM at pH 7.0, the latter increasing as the pH increases and becoming negligible at pH 8.0. The X-ray crystal structure of SPU inhibited by sulphite, determined at 1.65 Å resolution, provides a rationale for this kinetic behaviour. Indeed, the structure reveals that sulphite binds to the two Ni(II) ions of urease active site in a tridentate mode, by using all its three oxygen atoms. Such a binding mode is fully consistent with the competitive inhibition mode observed in the biochemical analysis. The analysis of the hydrogen-bonding network around the nickel-bound sulphite further provides a basis for the interpretation of the pH-dependence of inhibition.

The original papers describing the inhibition of urease by fluoride and sulphite are available in this dissertation as Annex 2 and Annex 3, respectively. The original citations are here below.

- Fluoride inhibition of *Sporosarcina pasteurii* urease: structure and thermodynamics.
  S. Benini, M. Cianci, L. Mazzei and S. Ciurli (2014), *J. Biol. Inorg. Chem.*, **19**:1243.

- Kinetic and structural studies reveal a unique binding mode of sulfite to the nickel center in urease.
  L. Mazzei, M. Cianci, S. Benini, L. Bertini, F. Musiani and S. Ciurli (2016), *J. Inorg. Biochem.*, **154**:42.

CHAPTER 3

CHEMISTRY AT THE UREASE SURFACE

3.1 A biochemical and structural study on two urease inactivators

After the two papers describing the inhibition of urease by fluoride and sulphite and their interactions within the active site, I have devoted my efforts in order to evaluate the possibility to modulate urease activity by blocking the conserved mobile flap that is involved in the regulation of substrate entrance into, and products exit from, the active site. Two papers resulted from this work, describing kinetic and structural studies on the inhibition of two urease inactivators (see APPENDIX A for further details regarding irreversible inhibition): 1) 1,4-Benzoquinone (BQ), and ii) catechol (CAT). For these studies, the bacterial urease from *Sporosarcina pasteurii* (SPU) and the plant urease from *Canavalia ensiformis* (jack bean, JBU), were used as model enzymes. Kinetic experiments were carried out by using a spectrophotometric method, while the structural characterization of the inhibition was achieved using X-ray crystallography.

The results of the kinetic study show that 1,4-Benzoquinone (BQ) irreversibly and stoichiometrically inhibits SPU, in a concentration- and time-dependent manner. The X-ray crystal structure of SPU inactivated by BQ, determined at 2.07 Å resolution, reveals the presence of a 1,4-hydroquinone moiety covalently bound to the thiol group of $Cys^{\alpha 322}$ that belongs to the mobile flap. This observation has been rationalized on the basis of the reactivity of thiols with benzoquinones, following a Michael-type reactivity.

In a similar fashion, the X-ray crystal structure of SPU inhibited by CAT, determined at 1.50 Å resolution, proved that catechol binds to the same thiol group of $Cys^{\alpha 322}$. In this

case, however, the kinetic data collected on the inhibition of JBU by catechol indicated the presence of a more complex inhibition mechanism, which involves an initial lag phase followed by a rapid decrease of urease activity, as a function of time. This inhibition profile was successfully modelled by assuming a radical-based autocatalytic multistep mechanism, the latter being proposed and discussed in the paper.

In general, BQ and CAT abolish urease activity through a stabilization of the flap in the open conformation. In the enzymatic mechanism of urease-catalysed urea hydrolysis, the flap must switch from an open to a closed conformation in order to stabilize the reaction intermediate. In the presence of BQ and CAT, the flap is blocked in the open conformation and the catalysis cannot occur.

Altogether, these results have provided strong evidences on the possibility to develop BQ and CAT derivatives able to interact with the cysteine residue and to act as urease inactivators. Both quinones and polyhydroxylated aromatic functionalities are found in a large number of natural compounds that can be potentially tested as urease inhibitors.

The original papers describing the inhibition of urease by BQ and CAT are available in this thesis as Annex 4 and Annex 5, respectively. The original citations are here below.

- Inactivation of urease by 1,4-benzoquinone: chemistry at the protein surface.
  L. Mazzei, M. Cianci, F. Musiani and S. Ciurli (2016), *Dalton Trans.*, **45**:5455.

- Inactivation of urease by catechol: Kinetics and structure.
  L. Mazzei, M. Cianci, F. Musiani, G. Lente, M. Palombo and S. Ciurli (2017), *J. Inorg. Biochem.*, **166**:182.

CHAPTER 4

UREASE INHIBITION BY NBPT,
A SUICIDE SUBSTRATE

4.1   Introduction

Phosphoramides are a class of well-known and very potent urease inhibitors, acting with a slow-binding inhibition mechanism.[1] The inhibition process involves a first hydrolytic event performed *in situ* by urease, with the subsequent formation of a tetrahedral moiety that blocks the enzyme active site by mimicking the transition state that would be formed during the enzyme reaction of urea hydrolysis. It has been demonstrated that, irrespective of the starting compound, the inhibition is always brought about by diamidophosphate (DAP).[2-4] In this context, significant information on the structural basis of urease inhibition by phosphoramides was provided by Benini and co-workers in 1999, who reported the X-ray crystal structure of the bacterial urease from *Sporosarcina pasteurii* (SPU) in complex with DAP after the treatment of the enzyme with phenylphosphorodiamidate (PPD).[5]

N-(n-butyl)thiophosphoric triamide (NBPT) is a particular derivative of such compounds, where an oxygen atom of the phosphoryl scaffold has been replaced by a sulphur atom. NBPT is the nitrogen stabilizer mostly used in agriculture to counteract ammonia volatilization resulting upon the use of urea as a nitrogen fertilizer and it is regarded as the most effective urease inhibitor in aerobic soils.[6] It is commercialized by *Koch Fertilizer, LLC*, under the commercial name of *Agrotain®* in a number of formulations. Several studies have been conducted on the efficacy of NBPT in inhibiting ureases from different sources and in reducing ammonia volatilization upon urea addition, either *in*

vitro or in soils.[7-12] These studies commonly assert that NBPT has little or no effect as a urease inhibitor, while it must be converted to its *oxo*-analogue, N-(n-butyl)phosphoric triamide (NBPTO), in order for a strong inhibition to occur. Also, the latter conversion takes place predominantly in soils, rather than in solution. In 2001, Kot and coworkers partially reverted previous evidences, demonstrating that NBPT itself is able to inhibit plant urease, even though its efficacy is smaller than its *oxo*-analogue.[13]

In the present work a biochemical and structural characterization of the NBPT inhibition on two ureases, the bacterial enzyme from *Sporosarcina pasteurii* (SPU) and the plant enzyme from *Canavalia ensiformis* (jack bean, JBU) is reported. The results of the kinetic experiments prove that NBPT acts as a slow-binding inhibitor of the enzyme. The X-ray crystal structure of SPU inhibited by NBPT, determined at 1.28 Å resolution, demonstrates that the inhibitor directly interacts with the nickel ions in the urease active site, undergoing an *in situ* hydrolysis that generates a tetrahedral moiety blocking the active site in a fashion similar to that proposed for phosphoramides.

4.2 Materials and methods

4.2.1 Enzyme and inhibitor sources

Urease from *Sporosarcina pasteurii* (SPU) was expressed and purified from the native source following a previously described procedure.[14] Protein quantification was carried out by measuring the activity using a pH-STAT method[15] and considering its specific activity of 2500 units mg^{-1} and M_r = 250 kDa. Urease from *Canavalia ensiformis* (jack bean, JBU) type C-3, powder (≥600,000 units/g) was purchased from *Sigma-Aldrich* and was quantified following manufacturer's information.

N-(n-butyl)thiophosphoric triamide was purchased from *Apollo Scientific*.

4.2.2 Enzymatic assays

Pre-incubation experiments were carried out at room temperature by using a spectrophotometric assay and an *Agilent Cary 60* UV-Vis Spectrophotometer. The pH indicator cresol red was used to monitor the overtime increase of pH due to urease

activity, following a previously described procedure.[14] In the case of SPU-NBPT experiments, a solution of 50 nM SPU dissolved in 2 mM HEPES buffer at pH 8.0, containing 1 mM Na_2SO_3, was diluted 50 times using a 2 mM HEPES buffer solution at pH 8.0, containing 30 mg L^{-1} cresol red and 2 mM EDTA (Buffer CR). In the case of JBU-NBPT experiments, a solution of 250 nM JBU dissolved in 20 mM HEPES buffer, at pH 8.0, was diluted 50 times in the same Buffer CR. Then, solutions containing different concentrations of NBPT, in 2 mM HEPES buffer at pH 8.0, were added to either the SPU or the JBU urease solutions prepared as described above (final concentration of NBPT in the range 0-400 µM). The time zero of the pre-incubation experiment was considered as the time of the mixing of the urease and the inhibitor solutions. Subsequently, at fixed times, aliquots of the pre-incubation solutions were taken and added to a solution containing urea to a final concentration of 100 mM, and the time-dependent absorbance change at 573 nm was monitored. The enzyme activity was calculated as the slope of the linear portion of the absorbance *vs.* time curve. The values of the residual activity at different pre-incubation time points were then calculated by normalizing each activity value to the activity measured at time zero of pre-incubation. Residual activities were finally expressed as percentage values, as a function of pre-incubation time.

4.2.3   Crystallization, data collection and structural determination

A 11 mg mL^{-1} urease solution in 20 mM HEPES buffer at pH 7.5, containing 50 mM Na_2SO_3 and 2 mM EDTA, was incubated for 1 h in the presence of increasing concentrations of NBPT (in the range 1-16 mM) dissolved in the same buffer. Subsequently, 2 µL of each SPU-NBPT solution was diluted with 2 µL of a precipitant solution consisting in 1.6-2.0 M ammonium sulphate dissolved in 50 mM sodium citrate buffer at pH 6.3, and containing the same concentration of NBPT. Crystallization trials were performed at 293 K using the hanging-drop method, equilibrating the drop against 1 mL of the precipitant solution using 24-well XRL Plates (*Molecular Dimensions*). Rice-shaped protein crystals appeared in most of the crystallization conditions within two weeks and grew to a size of 0.1 x 0.1 x 0.3 mm^3. Crystals were scooped up using cryoloops and transferred to a cryoprotectant solution of 20% ethylene glycol dissolved in

50 mM sodium citrate buffer at pH 6.3, also containing 2.4 M ammonium sulphate and the same concentration of NBPT present in the crystallization drop. The crystals were then flash-cooled and stored in liquid nitrogen.

Diffraction data were collected at 100 K using synchrotron radiation at the EMBL P13 beamline of the Petra III storage ring, c/o DESY, Hamburg (Germany). The beamline was equipped with a Si(111) crystal monochromator (*FMB Oxford*), a DECTRIS Pilatus 6M detector, and a MD2 goniometer (*MAATEL-EMBL*) with a horizontal spindle axis. Reflection images were recorded by performing helical scans along the crystal to achieve higher resolution by minimizing radiation damage.

Data processing and reduction was carried out with XDS[16] and AIMLESS.[17] The crystals were isomorphous with respect to those of native urease and other complexes of the same enzyme. The crystal structure of SPU in complex with diamidophosphate (PDB code 3UBP, 2.00 Å resolution)[5] devoid of solvent molecules and ligands, was used as a initial model for the rigid body refinement of the αβγ SPU trimer, carried out using Refmac.[18] Model building and water or ligand addition/inspection were conducted using Coot.[19, 20] The structure was isotropically refined, including the hydrogen atoms in the riding positions, and then anisotropically refined to final R and R_{free} of 11.18 and 13.99, respectively. The X-ray diffraction data and final refinement statistics are given in Table 1. Figures were generated using PyMol (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.), and CrystalMaker (http://www.crystalmaker.com).

Table 1: Data collection, processing and refinement statistics

Data collection	
Wavelength (Å)	0.9537
Detector	DECTRIS Pilatus 6M
Crystal-to-Detector distance (mm)	225.3
Oscillation angle (degrees)	0.100
Number of images	1300
Space group	$P6_322$
Unit cell (*a, b, c*, Å)	131.74, 131.74, 188.93
Resolution range (Å)[1]	97.67 – 1.28
Total number of reflections[1]	3479631 (171907)
Unique reflections[1]	243949 (11805)
Multiplicity[1]	14.3 (14.6)
Completeness[1] (%)	99.4 (98.1)
R_{sym}[1,2] (%)	0.117 (2.260)
R_{pim}[1,3] (%)	0.033 (0.625)
Mean I half-set correlation CC(1/2) [1]	0.999 (0.673)
Mean $I/\sigma(I)$[1]	15.9 (1.5)
Refinement statistics	
Number of monomers in the asymmetric unit	3
R_{factor}[4] (%)	11.18
R_{free}[4] (%)	13.99
Cruickshank's DPI for coordinate error[5] based on R_{factor} (Å)	0.030
Wilson plot B-factor ($Å^2$)	10.4
Average all atom B-factor[6] ($Å^2$)	16.645
RMS (bonds)[4]	0.01
RMS (angles)[4]	1.477
Total number of atoms	7055
Total number of water molecules	766
Solvent content (%)	55.52
Matthews Coefficient ($Å^3$/Da)	2.76
Ramachandran plot[7]	
Most favored regions (%)	89.9
Additionally allowed regions (%)	9.2
Generously allowed regions (%)	0.8
Disallowed regions (%)	0.2

[1]highest resolution bin in parentheses;

[2]$R_{sym} = \sum_{hkl}\sum_j |I_j - \langle I \rangle| / \sum_{hkl}\sum_j I_j$, where I is the intensity of a reflection, and $\langle I \rangle$ is the mean intensity of all symmetry related reflections j;

[3]$R_{p.i.m.} = \sum_{hkl}\{[1/(N-1)]\sum_j |I_j - \langle I \rangle|\}^{1/2} / \sum_{hkl}\sum_j I_j$, where I is the intensity of a reflection, and $\langle I \rangle$ is the mean intensity of all symmetry related reflections j, and N is the multiplicity;[21]

[4]taken from REFMAC;[18] R_{free} is calculated using 5% of the total reflections that were randomly selected and excluded from refinement;

[5]$DPI = R_{factor} \cdot D_{max} \cdot compl^{-1/3} \sqrt{\frac{N_{atoms}}{(N_{refl} - N_{params})}}$, where N_{atoms} is the number of the atoms included in the refinement, N_{refl} is the number of the reflections included in the refinement, D_{max} is the maximum resolution of reflections included in the refinement, *compl* is the completeness of the observed data, and for isotropic refinement, $N_{params} \approx 4N_{atoms}$;[22]

[6]taken from BAVERAGE;[23]

[7]taken from PROCHECK.[23]

4.3   Results and discussion

4.3.1   Biochemical characterization of NBPT inhibition on urease

The inhibition of NBPT on SPU and JBU was investigated by performing pre-incubation experiments at pH 8.0, and the results are shown in Figure 1 (see Appendix A for details about pre-incubation experiments). The calculated values for the percentage residual activity as a function of pre-incubation time were optimally fitted to a single exponential decay (Figures 1A and 1C). The resulting pseudo-first order constants k_{obs} (s^{-1}) showed a linear dependence on the concentration of NBPT (Figure 1B and 1D). The value for the second-order kinetic constant resulting from the linear fit of the k_{obs} vs. NBPT concentration plot is 29.9 ± 0.4 $M^{-1}s^{-1}$ and 24.2 ± 0.3 $M^{-1}s^{-1}$ for the inhibition of NBPT on SPU and JBU, respectively. This second-order kinetic constant refers to the association constant k_{on} that regulates the formation of the enzyme-inhibitor complex (see Appendix A for theoretical details regarding kinetic constants). The kinetic behaviour observed in the pre-incubation experiments is consistent with a slow-binding inhibition mechanism that follows a first-order kinetic with respect to both the two reactants (SPU/JBU and NBPT). Furthermore, the similar values obtained for k_{on} in the case of bacterial and plant ureases demonstrate that the reaction of NBPT with the active site of urease and the consequent formation of stabilizing interactions can be generalized among ureases from different organisms, without any species-specific effect. The measured values of k_{obs} (k_{on}) are similar but slightly different from the value previously reported by Kot and co-workers for the inhibition of JBU by NBPT (110 ± 10 $M^{-1}s^{-1}$).[13] Probably, this difference can be ascribed to the different experimental setup used. In fact, they have determined the association constant by performing time-course experiments. In such experiments, substrate (urea) and inhibitor are simultaneously added to the enzyme, so that competition effects can arise, affecting the measurement.

In order to determine the inhibition constant K_i, (given by the k_{off}/k_{on} ratio, see Appendix A for details) the value of the first-order rate constant k_{off}, that regulates the dissociation of the enzyme-inhibitor complex, would be also needed. In general, a good estimation of k_{off} can be derived from the y-intercept of the linear fit of the k_{obs} vs. [I] plot. As one can notice from Figures 1B and 1D, in the experimental conditions used for this measurement the linear fit appears to intercept at the origin, suggesting a very low dissociation constant

value. The value of k_{off} previously reported in the case of JBU inhibition by NBPT is $(0.17 \pm 0.03) \times 10^{-4}$ s^{-1},[13] confirming that NBPT hardly dissociates from the urease active site.
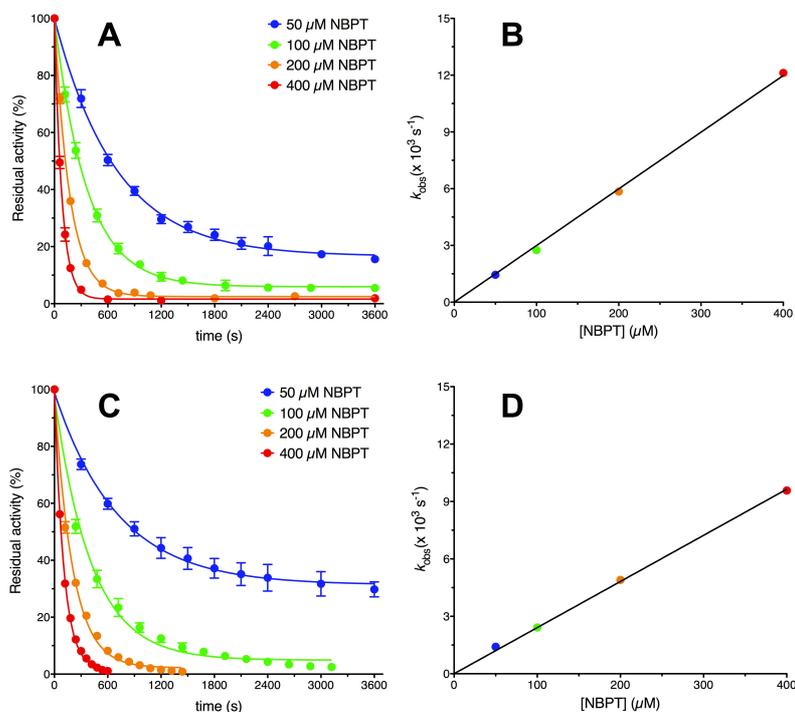
Figure 1: Inhibition of ureases by NBPT. (A and C) SPU and JBU residual activities *vs.* time plots at different NBPT concentrations, respectively. (B and D) Linear plots of k_{obs} as a function of NBPT concentration in the case of SPU and JBU, respectively. In both panels, the lines represent the result of an exponential or linear fit of the data.

4.3.2 Structural analysis of SPU inhibited by NBPT

The structure of SPU co-crystallized in the presence of NBPT reveals the well-known quaternary structure of *Sporosarcina pasteurii* urease, consisting of a $(\alpha\beta\gamma)_3$ trimer of trimers where the α subunit is formed by an $(\alpha\beta)_8$ barrel domain and a β-type domain, the β subunit is mainly characterized by β strands, and the γ subunit consists of $\alpha\beta$ domains. The refined crystallographic model closely matches that of native urease (PDB code 4CEU),[24] with a global root mean square deviation (RMSD) between their backbones equal to 0.17, 0.20 and 0.14 Å for the α, β and γ subunits, respectively. A more detailed per-residue analysis is presented in Figure 2. A comparison between the backbones of the

native and inhibited enzymes in terms of pairwise RMSD per residue (Figure 2A) displays a considerable superimposition of both the β and γ subunits. However, a substantial difference in the residue range 310-340 belonging to the α subunit comes to light. This region corresponds to a conserved highly flexible helix-loop-helix motif, also known as mobile flap, which controls the access of the substrate into the active site cavity of urease.[25] Several crystallographic studies on native SPU, as well as other inhibited forms of the enzyme, have described the flap in a open conformation.[25] A remarkable exception is given by the X-ray crystal structure of SPU inhibited by diamidophosphate (DAP), a transition state analogue that is generated *in situ* by the enzymatic hydrolysis of phenylphosphorodiamidate (PPD).[5] This model has the flap in a closed conformation, the latter being proposed to stabilize the intermediate of the catalysis. The evaluation of pairwise RMSD per residue between the backbones of the DAP- and NBPT-inhibited enzymes (Figure 2B) yields in small values for the region comprised in the 310-340 residue range of the α subunit, thus revealing the closed conformation of the mobile flap in the NBPT-inhibited SPU.
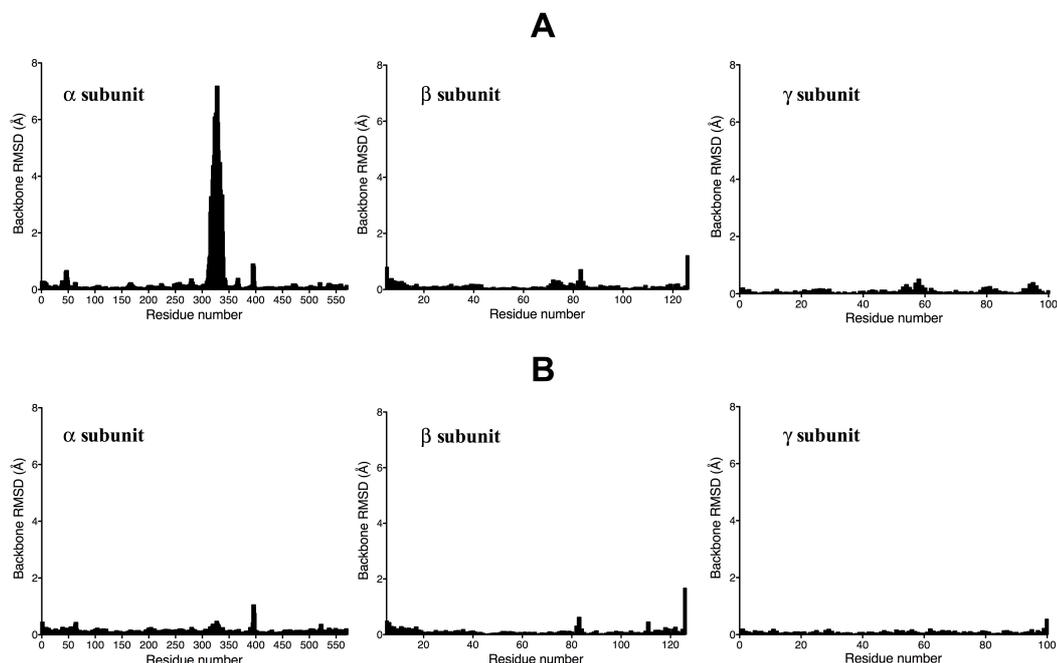
Figure 2: Pairwise root mean square deviation (RMSD) per residue. (A) Comparison between the α, β and γ subunits in the native (PDB code 4CEU) and NBPT-inhibited SPU. (B) Comparison between the α, β and γ subunits in the DAP- and NBPT-inhibited SPU. The region belonging to the mobile flap is comprised in the 310-340 residue range of the α subunit.

The electron density belonging to the protein active site is well defined (Figure 3). The overall architecture of the active site, in terms of those residues that directly interact with the two Ni(II) ions, is highly conserved with respect to the native enzyme.[5, 24] The two Ni(II) ions are well ordered [B-factors of 11.6 and 10.5 Å^2 for Ni(1) and Ni(2), respectively] and are separated by 3.7 Å, a distance similar to that found in the native enzyme (3.6 Å). The two Ni(II) ions are bridged by the carboxylate group of the carbamylated Lys$^{\alpha 220*}$, which is bound to Ni(1) by Oθ1 (at 2.0 Å) and to Ni(2) by Oθ2 (at 2.0 Å). Ni(1) is further coordinated by His$^{\alpha 249}$ Nδ (at 2.0 Å) and by His$^{\alpha 275}$ Nϵ (at 2.0 Å), whereas Ni(2) is bound to His$^{\alpha 137}$ Nϵ (at 2.1 Å), His$^{\alpha 139}$ Nϵ (at 2.1 Å) and Asp$^{\alpha 363}$ Oδ1(at 2.1 Å). The unbiased omit electron density map shown in orange in Figure 3 (calculated with Fourier coefficients F_o-F_c and phases from the refinement of NBPT-inhibited SPU structure using the DAP-inhibited SPU structure as an initial model after removal of the DAP moiety and other solvent/ligand molecules), as well as the $2F_o$-F_c electron density map, reveal the presence of additional electron density around the two Ni(II) ions that does not match the solvent molecules usually present in the active site of native SPU. This electron density has a tetrahedral shape and indicates the presence of a non-protein ligand that completes the coordination spheres of the two Ni(II) ions by binding to them through three atoms, with a fourth atom pointing away from the bimetallic centre. This tetrahedral arrangement exactly replaces the cluster of four water molecules existing in the active site of enzyme in its resting state.
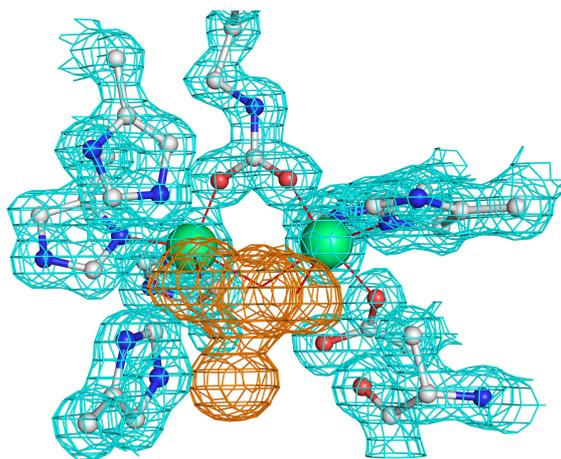
Figure 3: Atomic model of the active site of SPU inhibited by NBPT. The nickel-coordination environment is shown superimposed on the final $2F_o$–F_c electron density map contoured at 1.5 σ (cyan), while the unbiased F_o–F_c omit map corresponding to the ligand is shown contoured at 3 σ (orange). Carbon, nitrogen, oxygen and nickel atoms are grey, blue, red and green, respectively.

32

Considering the presence of NBPT in the crystallization solution, attempts to fit this entire molecule into such an electron density failed, because the latter lacks the portion ascribable to the NBPT butyl group. As described in the introduction of this chapter, the ability of urease to hydrolyse amides and esters of phosphoric acid *in situ* is well-documented in the literature.[2] Moreover, the crystallographic study on the inhibition of SPU by PPD (PDB code 3UBP), demonstrated that PPD undergoes an hydrolytic event by the enzyme to give phenol and DAP, the latter acting as the actual inhibitor by mimicking the tetrahedral transition state that would occur during the hydrolysis of urea.[5] NBPT is a thiophosphoric triamide, with an *n*-butyl amine as a group bound to the phosphorous atom. The enzymatic hydrolysis of such a compound would result in the formation of a diamidothiophosphate, with the release of *n*-butyl amine. It is also known that phosphoramides undergo uncatalysed hydrolysis in solution.[26] In the case of NBPT, such a hydrolysis would result in the loss of ammonia, rather than *n*-butyl amine (due to the smaller basicity of the former). Hence, in solutions of NBPT, three different species can simultaneously be present (Figure 4): i) N-(n-butyl)thiophosphoric triamide (NBPT), ii) N-(n-butyl)thiophosphoric diamide (NBPD), and iii) N-(n-butyl)thiophosphoric acid (NBPA), so that the three theoretical derivatives originated upon hydrolysis by urease would be i) di-amidothiophosphoric acid (DATP), ii) mono-amidothiophosphoric acid (MATP), and iii) thiophosphoric acid (TP), with their relative conjugate bases.
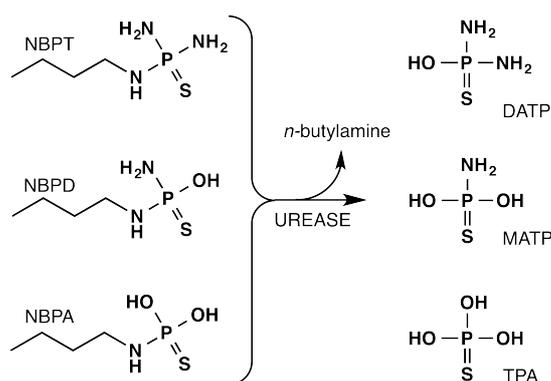
Figure 4: Schematization of the hydrolytic event catalysed by urease on NBPT derivatives. The three NBPT derivatives that can be originated by the uncatalysed reaction occurring in solution and the products generated by the subsequent enzyme hydrolysis are represented on the left and on the right, respectively. The hydrolysis performed by urease on all these derivatives results in the release of a molecule of *n*-butyl amine.

Three independent refinement procedures were therefore carried out by modelling DATP, MATP or TP into the additional electron density shown by the omit electron density map. The following analysis was thus performed in order to discern between the three possibilities on a structural basis and chemical grounds. Two main criteria were used to understand the correct orientation of the ligand bound to the two Ni(II) ions. First, the large scattering factor of sulphur allowed to unequivocally locate its position as the distal atom pointing away from the bimetallic centre. Second, as concluded in the case of the PPD-to-DAP conversion in the presence of SPU,[5] NBPT would undergo an initial nucleophilic attack on the P atom by the activated hydroxide ion located in the bridging position of the binuclear metallic centre, with the subsequent release of *n*-butyl amine and the formation of a tetrahedral moiety that has the hydroxyl functionality bridged between the two Ni(II) ions. These conclusions are schematically reported in Figure 5. In such a context, the identity of the terminal ligands L_1 and L_2 bound to Ni(1) and Ni(2), respectively, can be provided by a structural analysis, resulting in the identification of the actual moiety that binds to the active site of SPU. Indeed, as shown in Figure 5, the L_1 atom bound to Ni(1) must have two non-bonding electron pairs available, one deputed to coordinate Ni(1) and the other to receive a hydrogen bond from His$^{\alpha 222}$ Nε (at 2.71 Å) that is protonated, as inferred by the interaction of His$^{\alpha 222}$ Nδ with the peptide NH group of Asp$^{\alpha 224}$ (at 2.9 Å). In such a situation, an hypothetical P-NH$_2$ group would not satisfy this requirement, because in this case the only lone pair on the N atom would be needed to coordinate Ni(1); on the other hand, a P-OH group bound to Ni(1) would satisfy this criterion, because the O atom would have two lone pairs. A similar analysis revealed the presence of two hydrogen bonding acceptor atoms in the vicinity of the L_2 atom bound to Ni(2), namely the backbone carbonyl O atoms of Ala$^{\alpha 170}$ and Ala$^{\alpha 366}$ (both at 3.0 Å). This implies that L_2 acts as a two-hydrogen bonding donor, suggesting the presence of a P-NH$_2$ group bound to Ni(2). These preliminary structural analysis implies that MATP is the moiety bound in the active site of SPU, with two P-OH groups located at the bridging (L_B) and L_1 positions, respectively, and a P-NH$_2$ group located at the L_2 position, with the S atom located at the distant L_D position, pointing away from the Ni(II) ions and towards the active site entrance.
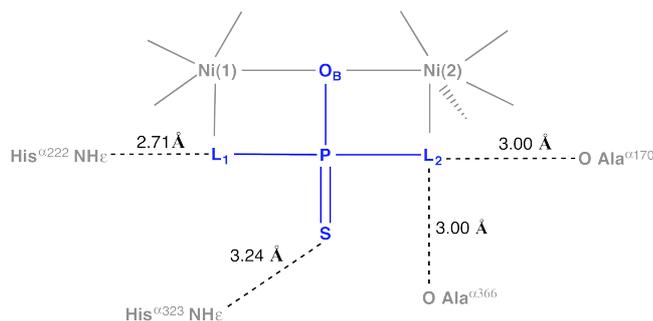
Figure 5: Structural scheme showing the hydrogen-bonding network between the nickel-bound thiophosphoramide derivative and the surrounding residues at the active site of SPU.

In order to support this conclusion, a statistical analysis of the B-factor values for the atoms bound to the Ni(II) ions in the alternative refinements was carried out (Table 2). The refined B-factor value provided for the L_1 atom bound to Ni(1) is significantly smaller in the case of DATP (9.6 $Å^2$) than in the case of MATP (12.1 $Å^2$) or TP (12.2 $Å^2$), and even smaller than that of the Ni(II) atoms, an unrealistic result that strongly supports the presence of a P-OH group in the L_1 position. The refined B-factor value provided for the L_2 atom bound to Ni(2) is significantly higher in the case of TP (14.3 $Å^2$) than that found for MATP (11.6 $Å^2$), the latter being very close to the B-factor of Ni(2) (10.5 $Å^2$), a more realistic result that strongly supports the presence of a $P-NH_2$ group at the L_2 position. Altogether, these structural data strongly support the presence of a MATP moiety as the actual inhibitor found in the active site of urease crystallized in the presence of NBPT.

Table 2: B-factor values for the independently refined DATP, MATP and TP at 1.28 Å

Ligand	DATP 1 @ 1.28 Å	MATP @ 1.28 Å	TP @ 1.28 Å
Ni(1)	11.5	11.5	11.5
Ni(2)	10.5	10.5	10.5
P	13.1	12.9	12.8
L_B	13.0	12.9	13.0
L_1	9.6	12.1	12.2
L_2	11.5	11.6	14.3
L_D	14.6	14.6	14.6

The final refinement procedure was therefore performed by modeling MATP in the active site of the inhibited urease crystal structure (Figure 6). MATP binds to Ni(1) and Ni(2) by its O_{L1} and N_{L2} atoms, respectively. The second MATP oxygen atom (O_B) symmetrically bridges the two Ni(II) ions, while the sulfur atom points away from the binuclear metallo-

center towards the active site cavity opening (Figure 6A). Refined selected distances and angles are given in Table 3. An extended network of hydrogen bonds stabilizes the MATP molecule within the active site cavity (Figure 6B). As previously described for the case of PPD-to-DAP,[5] the Ni(1)-bound MATP oxygen atom (O_{L1}) receives a hydrogen bond from His$^{\alpha 222}$ Nϵ (at 2.71 Å), while the Ni(2)-bound MATP NH$_2$ group (N_{L2}) acts as a donor of two hydrogen bonds to the carbonyl backbone O atoms of Ala$^{\alpha 170}$ (at 3.0 Å, not shown) and Ala$^{\alpha 366}$ (at 3.0 Å), respectively. Furthermore, the nickel-bridging MATP oxygen atom (O_B) is at hydrogen-bonding distance from Asp$^{\alpha 363}$ Oδ2 atom (2.6 Å), implying the presence of a proton shared between the latter and the nickel-bridging MATP oxygen. Finally, the sulfur atom of MATP is at hydrogen-bond distance from the His$^{\alpha 323}$ Nϵ (3.24 Å). The P-S bond length (1.95 Å) in MATP molecule is consistent with the presence of a double bond between the phosphorous and the sulfur atoms, thus suggesting the protonated state of His$^{\alpha 323}$ Nϵ. Moreover, it strongly supports the idea that the actual inhibitor is the neutral amidothiophosphoric acid (NH$_2$)PS(OH)$_2$, with the orientation of MATP in the active site of urease that closely matches that of DAP in the structure of SPU inhibited by DAP.[5]

Table 3: Selected distances and angles around the Ni(II) ions in the crystal structure of native SPU (PDB code 4CEU), DAP-inhibited SPU (PDB code 3UBP) and MATP-inhibited SPU

Ni - L Distances (Å)	4CEU @ 1.50 Å	3UBP @ 2.00 Å	SPU-MATP @ 1.28 Å
Ni(1) - Lys220* Oθ1	1.94	2.06	1.98
Ni(1) - L$_B$	2.08	2.32	2.16
Ni(1) - L$_1$	2.24	2.24	2.12
Ni(1) - His249 Nδ	2.03	1.99	2.00
Ni(1) - His275 Nε	2.02	2.10	2.02
Ni(2) - Lys220* Oθ2	2.08	1.92	2.05
Ni(2) - L$_B$	2.12	2.34	2.22
Ni(2) - L$_2$	2.07	2.32	2.17
Ni(2) - His137 Nε	2.11	2.13	2.10
Ni(2) - His139 Nε	2.08	2.16	2.11
Ni(2) - Asp363 Oδ1	2.10	2.12	2.09
Ni(1) ••• Ni(2)	3.67	3.82	3.72
L$_1$ ••• L$_2$	2.37	2.55	2.60
L - Ni - L Angles (°)			
Lys220* Oθ1 - Ni(1) - His249 Nδ	100.4	111.0	101.5
Lys220* Oθ1 - Ni(1) - His275 Nε	107.2	105.6	104.0
Lys220* Oθ1 - Ni(1) - L$_B$	96.6	95.4	96.8
Lys220* Oθ1 - Ni(1) - L$_1$	108.2	102.7	105.3
His249 Nδ - Ni(1) - His275 Nε	98.6	103.0	99.0
His275 Nε - Ni(1) - L$_B$	94.6	87.4	94.9
L$_B$ - Ni(1) - L$_1$	67.0	66.9	69.4
L$_1$ - Ni(1) - His249 Nδ	89.3	87.9	87.2
His249 Nδ - Ni(1) - L$_B$	154.2	147.2	153.5
His275 Nε - Ni(1) - L$_1$	141.6	143.2	148.0
Lys220* Oθ2 - Ni(2) - His137 Nε	90.8	90.0	90.4
Lys220* Oθ2 - Ni(2) - His139 Nε	91.7	94.0	91.5
Lys220* Oθ2 - Ni(2) - L$_2$	92.9	99.4	98.7
Lys220* Oθ2 - Ni(2) - L$_B$	95.6	99.8	96.9
Asp363 Oδ1 - Ni(2) - His137 Nε	82.8	79.7	81.5
Asp363 Oδ1 - Ni(2) - His139 Nε	86.4	82.4	83.4
Asp363 Oδ1 - Ni(2) - L(2)	94.5	92.3	91.4
Asp363 Oδ1 - Ni (2) - L$_B$	89.1	88.0	91.6
L$_2$ - Ni(2) – L$_B$	67.7	70.6	69.3
L$_B$ - Ni(2) - His137 Nε	95.0	92.6	93.4
His137 Nε - Ni(2) - His139 Nε	108.5	107.9	107.6
His139 Nε - Ni(2) - L(2)	88.4	86.9	88.7
Lys220* Oθ2 - Ni(2) - Asp363 Oδ1	172.4	167.5	168.6
L$_B$ - Ni(2) - His139 Nε	155.3	155.2	157.4
His137 Nε - Ni(2) - L$_2$	162.6	161.9	161.2
Ni(1) - L$_B$ - Ni(2)	122.1	110.2	116.6

As described above, the conformational change involving the closing (or the opening) of the flap does not significantly modify the position of the active site residues directly involved in the binding of the Ni(II) ions. However, this event drastically changes the positions of some of the amino acid side chains that, though not directly involved in nickel binding, face the active-site cavity and are considered important for the catalytic mechanism. In particular $Cys^{\alpha322}$, $His^{\alpha323}$ and $His^{\alpha324}$, which belong to the flap region, are approximately shifted by 5 Å with respect to the structure of native SPU. The backbone of the $Ala^{\alpha366}$ residue is also affected by the conformational change of the flap, even though it does not belong to the flap region: while in the native SPU the carbonyl O atom of $Ala^{\alpha366}$ is turned away from the Ni(II) ions and points towards the active site entrance, in the SPU-MATP structure it is turned towards the bimetallic active site, at hydrogen-bonding distance from the MATP nitrogen atom bound to Ni(2).
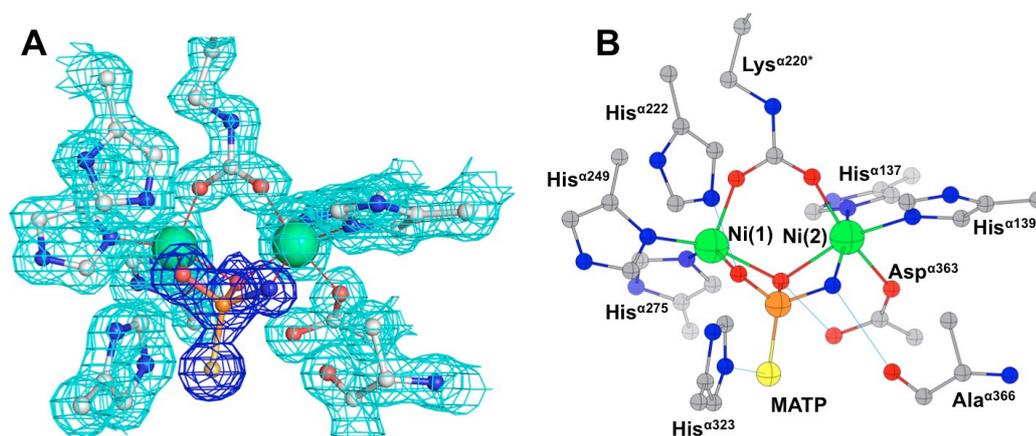
Figure 6: Atomic model of the active site of MATP-bound SPU. (A) Nickel-coordination environment is shown superimposed on the final $2F_o–F_c$ electron density map contoured at 1.5 σ. The map of the inhibitor is shown in blue. **(B)** The crystallographic structure of the same environment is represented. Putative hydrogen bonds are shown as thin blue lines. Spheres are drawn using the relative atomic radii values in CrystalMaker. Carbon, nitrogen, oxygen, sulphur, phosphorous and nickel atoms are grey, blue, red, yellow, orange and green, respectively.

These structural results further corroborate the current most accepted model for the reaction mechanism through which urease catalyses the hydrolysis of urea (described in details in Annex 1). The closing of the flap and the peculiar tetrahedral shape of MATP, as well as DAP, induce the formation of a strong hydrogen-bond network that blocks the flap in a closed conformation, thus disabling the protein for further substrate hydrolysis.

4.4   Conclusions and future perspectives

In this work a biochemical and structural study on the inhibition of urease by N-(n-butyl)thiophosphoric triamide (NBPT) has been carried out. Kinetic experiments have pointed out the slow-binding inhibition mode of NBPT. Structural data support the idea that NBPT in its mono-deaminated form first interacts with the active site. Our current hypothesis is that, independently on the rate of the uncatalysed hydrolysis of NBPT to give NBPD or NBPA and the concomitant presence of other hydrolysed species, the active site of urease will select only NBPD to enter the active site, the latter being the only NBPT derivative presenting the correct functionalities in order for its interaction with the bimetallic active site to be stabilized and the following hydrolytic event to occur. After the binding to the two Ni(II) ions, NBPD undergoes *in situ* hydrolysis with the formation of a tetrahedral moiety that mimics the transition state of the reaction of urea hydrolysis. This intermediate analogue stabilizes the mobile flap, involved in the entrance of substrate in the active site, in a closed conformation, thus precluding the protein from further substrate hydrolysis.

It is worth noting that this work is not to be understood as complete as some issues are still to be clarified. For example, efforts will be spent with the aim to research experimental evidences other than those provided by crystallographic results that would corroborate our structural information on MATP as the actual ligand in the active site of urease. These studies will involve both quantum-mechanical and/or docking calculations and analytic experiments, such as high-resolution mass spectrometry and ^{31}P and ^{1}H NMR. Another aspect to deal with is the determination of a reliable inhibition constant, K_i. First, differently from inhibitors that interact with enzymes through a rapid equilibrium, inhibition of slow-binding inhibitors is more complex to address because of their intrinsic time-dependent behaviour. Also, our data are consistent with the description of NBPT as a competitive inhibitor of urease, competing with urea for the binding to the active site. This fact complicates matters further, because the measured K_i value may also be affected by the concentration of substrate used in the experiments.

Undeniably, this is the first study showing on a structural basis the efficacy of NBPT in inhibiting urease, a role so far ascribed only to NBPTO, the NBPT *oxo*-analogue. Globally, the biochemical and structural details provided by this work, together with the

future results that will clarify the aspects described above, will pave the way for a structure-based design with the aim to develop more efficient urease inhibitors necessary to regulate its activity and to counterbalance its adverse effects.

REFERENCES

1. B. Krajewska (2009), *J. Mol. Catal. B Enzym.*, **59**:9.
2. R. Andrews, A. Dexter, R. Blakeley and B. Zerner (1986), *J. Am. Chem. Soc.*, **108**:7124.
3. G. W. McCarty, J. M. Bremner and J. S. Lee (1990), *Plant and Soil*, **127**:269.
4. W. S. Faraci, B. V. Yang, D. O'Rourke and R. W. Spencer (1995), *Bioorg. Med. Chem.*, **3**:605.
5. S. Benini, W. R. Rypniewski, K. S. Wilson, S. Miletti, S. Ciurli and S. Mangani (1999), *Structure*, **7**:205.
6. C. J. Watson (2000), International Fertiliser Society, York, UK, pp. 1.
7. G. W. McCarty and J. M. Bremner (1989), *Biol. Fertil. Soils*, **8**:290.
8. G. W. McCarty, J. M. Bremner and H. S. Chai (1989), *Biol. Fertil. Soils*, **8**:123.
9. C. B. Christianson, B. H. Byrnes and G. Carmona (1990), *Fertil. Res.*, **26**:21.
10. G. L. Creason, M. R. Schmitt, E. A. Douglass and L. L. Hendrickson (1990), *Soil Biol. Biochem.*, **22**:209.
11. L. L. Hendrickson and E. A. Douglass (1993), *Soil Biol. Biochem.*, **25**:1613.
12. D. G. Keerthisinghe and R. L. Blakeley (1995), *Soil Biol. Biochem.*, **27**:739.
13. M. Kot, W. Zaborska and K. Orlinska (2001), *J. Enzyme Inhib.*, **16**:507.
14. L. Mazzei, M. Cianci, S. Benini, L. Bertini, F. Musiani and S. Ciurli (2016), *J. Inorg. Biochem.*, **154**:42.
15. S. Ciurli, C. Marzadori, S. Benini, S. Deiana and C. Gessa (1996), *Soil Biol. Biochem.*, **28**:811.
16. W. Kabsch (2010), *Acta Crystallogr. D*, **66**:125.
17. P. Evans (2006), *Acta Crystallogr. D*, **62**:72.
18. G. N. Murshudov, A. A. Vagin and E. J. Dodson (1997), *Acta Crystallogr. D*, **53**:240.
19. P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan (2010), *Acta Crystallogr. D*, **66**:486.
20. P. Emsley and K. Cowtan (2004), *Acta Crystallogr. D* **60**:2126.
21. M. Weiss (2001), *J. Appl. Crystal.*, **34**:130.
22. D. W. Cruickshank (1999), *Acta Crystallogr. D*, **55**:583.
23. M. D. Winn, et al. (2011), *Acta Crystallogr. D*, **67**:235.
24. S. Benini, M. Cianci, L. Mazzei and S. Ciurli (2014), *J. Biol. Inorg. Chem.*, **19**:1243.
25. M. J. Maroney and S. Ciurli (2014), *Chem. Rev.*, **114**:4206.
26. J. D. Chanley and E. Feageson (1963), *J. Am. Chem. Soc.*, **85**:1181.

CONCLUSIONS

Urease as is a nickel-dependent enzyme that catalyses urea hydrolysis, triggering an overall pH increase and causing negative effects for human health as well as agriculture. Therefore, a tight control of its activity is required and several classes of inhibitors were studied in the last decades.

In this thesis, kinetic and structural studies of urease inhibition by a list of known inhibitors have been presented. The enzyme used for most of the experiments has been urease from the soil bacterium *Sporosarcina pasteurii* (SPU). In some cases, as the active site and, more in general, the overall structure of ureases is conserved, a plant urease from *Canavalia ensiformis* (jack bean, JBU) has been chosen as an alternative or a complement. The two main experimental approaches used in this work have been kinetic experiments and X-ray crystallography. In particular, five chemical species have been characterized both kinetically and structurally regarding their inhibitory properties towards urease: i) fluoride, ii) sulphite, iii) 1,4-benzoquinone (BQ), iv) catechol (CAT), and v) N-(n-butyl)thiophosphoric triamide (NBPT). The kinetic characterization of SPU inhibition by fluoride and sulphite has been carried out through isothermal titration calorimetry. Both these ligands act as reversible and pH-dependent inhibitors. In particular, fluoride acts with a mixed inhibition mechanism, while sulphite is a competitive inhibitor. The crystal structures of SPU inhibited by fluoride and sulphite, determined at 1.59 Å and 1.65 Å, respectively, have revealed that both inhibitors directly interact with the two Ni(II) ions in the active site. The study on the inhibition of urease by fluoride allowed an accurate depiction of the rationale for the inhibition mechanism, which further corroborated the currently accepted model for the catalytic mechanism. On the other hand, the coordination of sulphite to the two Ni(II) ions in the active site is unprecedented, and could stimulate scientists to develop inhibitors based on this scaffold. Differently from the previous cases, BQ and CAT have been kinetically characterized by using a spectrophotometric method. They both act as irreversible inhibitors of urease. The

inhibition mechanism of BQ, determined on SPU as a model enzyme, displays a classical exponential decay profile, with a very potent binding efficacy. Differently from the former, CAT has revealed a more complex inhibition behaviour, with an initial lag phase followed by a rapid increase of inhibition potency over time that has been interpreted as an inactivation process in which an autocatalytic radical-based mechanism is involved. X-ray crystal structures of SPU bound to BQ and CAT (determined at 2.07 Å and 1.50 Å resolution, respectively) have revealed that they covalently bind to a conserved cysteine residue located on a flexible flap that, even though not directly involved in the catalysis, modulates the access of the substrate urea into the active site cavity. So far, these X-ray crystal structures are the only two protein structures deposited in the PDB showing a covalent inhibition caused by a quinone or catechol as an extra-protein ligand.

In the last part of this work, a biochemical and structural characterization of urease inhibition by NBPT, a commercial product extensively used as a nitrogen stabilizer in agriculture, has been carried out. Biochemical experiments, performed by using a spectrophotometric method, have demonstrated that NBPT acts as a slow-binding inhibitor of urease, with an inhibition efficacy similar for bacterial and plant ureases. The X-ray crystal structure of SPU inhibited by NBPT disclosed that NBPT acts a suicide substrate. In fact, it is enzymatically hydrolysed *in situ*, with the formation of a tetrahedral product that directly interacts with the two Ni(II) ions in the active site with a very small dissociation constant. This is the first study that demonstrates, on a structural point of view and at atomic level, that NBPT is able to inhibit urease, reverting previously literature data that ascribed the inhibition process to its *oxo*-analogue, N-(n-butyl)phosphoric triamide (NBPTO).

The data reported in this dissertation, together with the knowledge on urease inhibitors already available in the literature, will pave the way towards a rational structure-based drug design, with the aim to develop new molecules acting as urease inhibitors. In particular, the three main scaffolds that will be investigated are quinones, polyphenols and phosphoramides, the latter in both their *thio-* and *oxo*-form. The analysis of the structure-activity relationships of the modified scaffolds will be of primary importance to further guide the drug discovery process, with the main goal being the development of new molecules with potential applications in medicine as well as in agro-environmental settings.

APPENDIX A

AN OUTLINE OF BASIC ENZYMOLOGY

A.1   A historical recap of enzymology

Enzymology refers to a branch of science concerned with the study of enzymes. Historically, the term enzyme was first provided in 1877 by the German scientist Kühne, who combined concepts that the French chemists Payen and Pasteur previously learned by studying intra- and extra-cellular alcoholic fermentation in yeasts. In that period, biochemical identity of enzymes was unknown and "vitalism" (the hypothesis that living and non-living organisms are distinct because the former contain some non-physical element and are governed by different non-physical principles than the latter) was still a more than accepted scientific hypothesis. In the following years, two important milestones contributed to inaugurate enzymology, as we know it today. The first was provided by Buchner's experiments in 1897, through which he demonstrated the ability of yeast extracts to carry out sugar fermentation processes even in the absence of living whole yeast cells.[1] The second milestone was achieved in 1926, when James Sumner reported that enzymes were proteins and could be crystallized.[2] For these extraordinary discoveries Buchner and Sumner were awarded the Nobel Prize in Chemistry in 1907 and 1946, respectively. Looking more deeply into the enzyme mode of action, the initial evidence that enzymes form a complex with substrates in order for the reaction to occur was given in 1880 by Wurtz[3] and in 1902 by Brown,[4] who pointed out the peculiar behaviour of the interaction between enzymes and substrates that leads to the enzyme saturation effect. The qualitative evidences were finally rigorously defined in mathematical form by Henry,[5] Michaelis and Menten,[6] and Briggs and Haldane,[7], who gave a great contribution to modern biochemistry. The brief list stated above is only an

example on how, in *ca.* fifty years, science went through the belief that living organisms were driven by vitalistic forces to a thorough understanding of the physical principles that regulate chemical reactions in living organisms, reaching a more chemical and molecular awareness. In the next paragraphs a brief description of the kinetics involved in enzyme-catalysed reactions will be provided, focusing on the role of inhibitors in modulating enzyme activity.

A.2   Enzyme kinetics

An enzyme (E in the following schemes and equations) is a biological macromolecule, in most cases a protein, able to catalyse a certain chemical reaction in which a substrate (S) is transformed into a product (P), as described in the simplified Scheme 1. In order for this process to occur, enzyme and substrate must first interact in order to form a complex (ES) in the active site of the enzyme.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \overset{k_2}{\longrightarrow} E + P$$
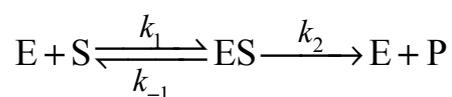
Scheme 1

ES complex represents a thermodynamic equilibrium, the latter being governed by k_1, namely the second-order rate constant (M^{-1} s^{-1}) that regulates the association rate of substrate end enzyme to give the ES complex, and by k_{-1}, namely the first-order rate constant (s^{-1}) that regulates the dissociation rate of the complex. The equilibrium dissociation constant (K_S) that numerically quantifies this process is described by the ratio of k_{-1} to k_1:

$$K_S = \frac{[E][S]}{[ES]} = \frac{k_{-1}}{k_1} \tag{A.1}$$

Where [E], [S] and [ES] are the molar concentrations of E, S and ES, respectively. According to Scheme 1, the overall velocity (*v*) of an enzyme-catalysed reaction is

directly proportional to the concentration of the ES complex, and it can be described by the following first-order equation:

$$v = k_2[\text{ES}] \tag{A.2}$$

A first mathematical description of such a proportionality was developed by Henry[5] and by Michaelis and Menten[6] in 1903 and 1913, respectively. The so-called rapid-equilibrium model assumes that, at any time, the concentration of substrate is much larger than the concentration of enzyme, the latter acting as a catalyser. Moreover, the velocity is measured in the initial stages of the reaction (initial velocity), so that the concentration of substrate already converted into products is negligible. In such an approach, rapid equilibrium between enzyme and substrate to form ES complex is followed by a slower conversion of ES complex to free enzyme and product(s), resulting in $k_2 \ll k_{-1}$, as defined in Scheme 1. According to this model, the ES complex can be described by the following expression:

$$[\text{ES}] = \frac{[\text{E}][\text{S}]}{K_\text{S} + [\text{S}]} \tag{A.3}$$

Where K_S is the equilibrium dissociation constant as defined in Equation A.1. After the formation of the ES complex, the reaction proceeds through a number of subsequent chemical steps to give product(s). If the conversion passes through a single chemical step, the process is governed by the first-order rate constant k_2, as shown in Scheme 1. In most of cases, however, product formation involves a series of rapid chemical events following the formation of ES complex. In such a complex contest, the rapid-equilibrium model includes all the rate constants controlling each singular chemical step in an overall first-order catalytic rate constant k_cat, with the following rearrangement of Scheme 1:

$$\text{E} + \text{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \text{ES} \xrightarrow{k_\text{cat}} \text{E} + \text{P}$$
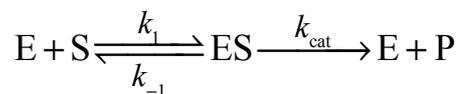
Scheme 2

So that in the expression for the reaction velocity (Equation A.2) k_2 is replaced by k_{cat} as follows:

$$v = k_{cat}[ES]$$ (A.4)

The final expression for the rapid-equilibrium model, referred to as Henry-Michaelis-Menten equation, combines Equations A.3 and A.4, in order to describe the dependence of initial velocity as a function of [S]:

$$v = \frac{k_{cat}[E][S]}{K_S + [S]}$$ (A.5)

It has to be noted that at infinite substrate concentration, all the molecules of enzyme will be saturated by the substrate, so that [ES] = [E]. Therefore, at very high substrate concentration the term $k_{cat}[E]$ tends towards a maximum value, referred to as the maximum reaction velocity V_{max}:

$$V_{max} = k_{cat}[E]$$ (A.6)

An implementation to the rapid-equilibrium model was provided in 1925 by Briggs and Haldane, with the steady-state model.[7] This model does not need that a rapid equilibrium occurs (k_2 << k_{-1} is not a required assumption anymore), rather it assumes that the reaction will go towards a steady-state phase, during which the concentration of ES complex is constant, being the rate of ES formation (governed by k_1) balanced by the dissociation rate of ES, either to give products or to dissociate back in E + S, and both giving the free form of the enzyme (governed by k_{-1} and k_2). In steady-state conditions, the concentration of the ES complex as a function of [S] can be described by the following expression:

$$[ES] = \frac{[E][S]}{[S] + \frac{\left(k_{-1} + k_2\right)}{k_1}} \tag{A.7}$$

As for the rapid-equilibrium model, we can assume the catalytic rate constant k_{cat} as the overall rate constant that takes into account the contribution of singular rate events following the ES complex formation, so that we can combine Equations A.4 and A.7 to obtain the expression for the initial velocity as a function of [S]:

$$v = k_{cat}[E] \frac{[S]}{[S] + \frac{\left(k_{-1} + k_2\right)}{k_1}} \tag{A.8}$$

If we set the following simplification for the denominator:

$$\frac{(k_{-1} + k_2)}{k_1} = K_M \tag{A.9}$$

Equations A.6, A.8 and A.9 can be combined to give the final expression for the steady-state model:

$$v = \frac{V_{max}[S]}{[S] + K_M} \tag{A.10}$$

Where K_M is commonly referred to as Michaelis constant. Even though the above mathematical derivation has been provided by Briggs and Haldane[7], Equation A.10 is commonly known as the Henry-Michaelis-Menten equation.

A comparison between the expressions for the rapid-equilibrium and for the steady-state models allows one to realize that they slightly differ one to the other. In particular, in the rapid-equilibrium model the equilibrium dissociation constant K_S is used, while the

steady-state model replaces it with K_M. As stated above, K_s is a thermodynamic constant represented by the ratio of k_{-1} to k_1, whereas K_M has to be treated as a kinetic constant, rather than a thermodynamic one, represented by the ratio of $(k_{-1} + k_2)$ to k_1. Therefore, despite different physical properties, K_s and K_M become equivalent under the specific conditions in which $k_2 << k_{-1}$.

A.2.1 Biochemical significance of K_M and k_{cat}

On a chemical point of view, K_M is a combination of first-order rate constants (k_{-1} and k_2) with a second-order rate constant (k_1). Therefore, it has units of molarity, as well as [S]. A first important biochemical significance of K_M can be highlighted by carrying out an experimental measurement of the initial velocity at a concentration of substrate that exactly matches K_M. In such conditions, Equation A.10 can be rewritten as follows:

$$v = \frac{V_{max}[S]}{[S]+[S]} = \frac{V_{max}}{2} \tag{A.11}$$

From this treatment we can evince that K_M experimentally matches the concentration of substrate yielding a half-maximal reaction velocity in the experimental conditions used. K_M is a specific parameter characteristic of any enzyme-substrate pair. Even though it is not an equilibrium dissociation constant, K_M directly reflects the reciprocal binding affinity for a certain enzyme towards a substrate. In general, for an enzyme able to catalyse a certain reaction on more than one substrate individually, a different K_M value is reported for each enzyme-substrate couple. Urease, for instance, can hydrolyse urea and hydroxyurea among other substrates, with the two K_M values falling in the range 1 - 4 mM and 1 - 125 mM, respectively.[8] Moreover, either by changing the experimental conditions, such as pH, temperature or ionic strength, or by using mutant species of an enzyme (with mutations belonging to residues involved in the enzyme-substrate recognition), K_M values for an enzyme-substrate pair may be strongly affected.

As described above, k_{cat} is defined as a first-order catalytic constant that regulates the overall rate velocity of an enzyme-catalysed reaction. In literature, k_{cat} is also known as the turnover number for the enzyme (s^{-1}) in the presence of a certain substrate, describing

the number of catalytic events that take place per unit time. Biochemically, it represents the maximum rate of an enzyme reaction in the presence of infinite substrate availability. In particular, it is related to the biochemical steps subsequent to the formation of ES complex, so that alterations in the late phases of enzyme reaction would yield in changes of k_{cat}. Similarly to K_M, k_{cat} perturbations can be caused by enzyme mutations as well as specific parameters related to the experimental conditions used. Since k_{cat} is usually the result of the multiple contribution of sequential biochemical steps, it provides a lower edge on the first-order rate constant of the slowest step following substrate binding, i.e. the rate-determining step, whereas it can not give information on the rates of singular steps.

As a conclusive remark, it has to be noted that the best description of the catalytic efficiency of a certain enzyme for a substrate is given by the ratio of k_{cat} over K_M (M^{-1} s^{-1}). This is because substrate specificity often results from differences in the transition state, which is an activated E*S complex that it is formed after the ES complex formation. Therefore this would give differences in k_{cat} rather than in K_M, the latter being affected by the formation of initial ES complex only.

A.2.2 Experimental determination of the kinetic parameters

Nowadays, the most common way to determine the kinetic constants k_{cat} and K_M for a certain enzyme reaction consists in measuring the initial velocity at increasing concentrations of substrate (through one of the several experimental techniques available[9]), as shown in Figure 1. The obtaining results can be plotted and directly fitted by using Equation A.10 (Figure 1A).[10] However, the widespread availability of user-friendly nonlinear curve-fitting programs is relatively recent. In the past, determination of the kinetic constants was carried out by transforming the data to produce linear plots. The most commonly used method for linearizing such experimental data is the Lineweaver and Burk approach.[11] Starting from the Michaelis-Menten equation, some rearrangements can be done in order to obtain a reciprocal form of the same equation, as follows:

$$\frac{1}{v} = \left(\frac{K_{\mathrm{M}}}{V_{\max}} \frac{1}{[\mathrm{S}]} \right) + \frac{1}{V_{\max}} \qquad\qquad (\mathrm{A.12})$$

This expression is described as a classical equation for a straight line (y= ax + b) where a is the slope and b is the y-intercept, while the independent and dependent variables x and y are described by the reciprocal value for the substrate concentration and the initial velocity, respectively (Figure 1B).[10] In such a representation, the values for K_{M} and $V_{\max}$ can be graphically determined. In particular, the y-intercept corresponds to $1/V_{\max}$ value, while the slope of the resulting linear fit defines $K_{\mathrm{M}}/V_{\max}$. Additionally, the x-intercept value defines the $-1/K_{\mathrm{M}}$ value. Even though the Lineweaver-Burk linearization approach is out-dated for enzyme-substrate kinetic parameters determination, it is still a valuable tool for certain purposes, as enzyme inhibition studies. Use of this approach will be discussed in the next paragraph.
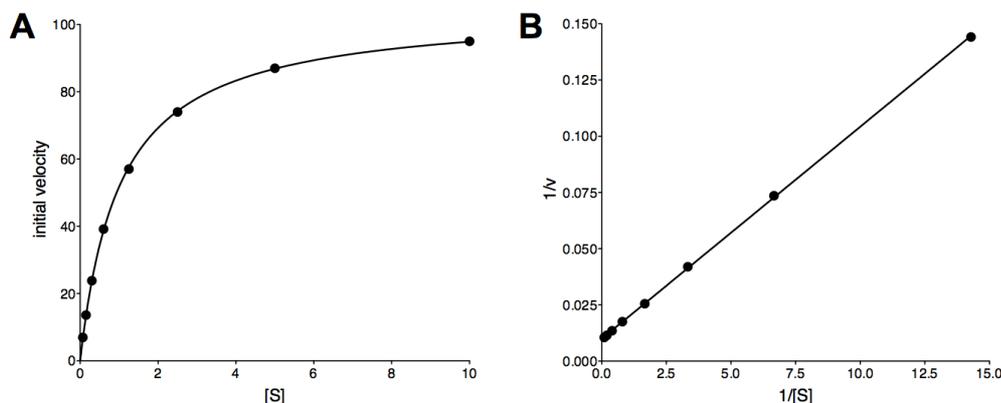
Figure 1: Plots of initial velocity as a function of substrate concentration. (A) Initial velocity measurements of a typical enzyme carried out at increasing concentrations of substrate. The line corresponds to the fit of the data performed according to Equation A.10. (B) Lineweaver-Burk linearization of data in (A). In this case, the line corresponds to a linear fit of the data. Redrawn from the book chapter: *Enzyme reaction mechanisms*.[10]

A.3 Enzyme inhibition

Enzyme inhibition is a biochemical process through which the rate of an enzyme-catalysed reaction is reduced or totally abolished by a ligand, namely an inhibitor. Modern chemists and biochemists spend efforts in studying the inhibition of enzymes for a number of reasons, most of them being related to pharmaceutical and agricultural

aspects. First, inhibitors are used as drugs to counteract diseases and metabolic dysfunctions caused by enzyme alterations.[12] Also, inhibitors are commonly used in agriculture to contrast the negative effects potentially caused by insects and parasites, as well as soil processes that would determine problems in plant growth.[13] Inhibitors are commonly sorted in two major classes, according to their inhibition mode: i) reversible inhibitors, that bind the enzyme in a reversible manner, and ii) irreversible inhibitors, also named inactivators, which irreversibly block the enzyme upon the formation of a covalent bond between some critical functionalities of the enzyme and the inhibitor itself.

A.3.1 Reversible inhibitors

In general, a reversible inhibitor (I) reversibly binds to a target enzyme to give an enzyme-inhibitor binary complex (EI) according to the following simplified reaction scheme:

$$E + I \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} EI$$
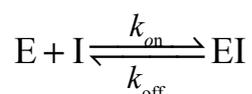
Scheme 3

This chemical process is governed by k_{on}, namely the second-order rate constant ($M^{-1} s^{-1}$) that regulates the association of inhibitor end enzyme to give the EI complex, and by k_{off}, namely the first-order rate constant (s^{-1}) that regulates the dissociation of the EI complex. As for the case of the ES complex formation, a thermodynamic equilibrium can be quantified by the dissociation constant K_i (M), the latter being defined by the ratio of k_{off} to k_{on}:

$$K_i = \frac{k_{off}}{k_{on}} \tag{A.13}$$

Where K_i also quantifies the inhibition strength of the inhibitor on the enzyme molecule, by means that at an inhibitor concentration equal to K_i value the enzyme is half-saturated by the inhibitor itself. To better understand the chemical bases of reversible inhibition, we should take into account that, in solution, the global system consists of three components,

the enzyme, the substrate and the inhibitor. Therefore, equilibrium is established as follows:

$$\begin{array}{ccc}
E + S & \underset{K_s}{\overset{}{\rightleftharpoons}} & ES \xrightarrow{k_{cat}} E + P \\
+ & & + \\
I & & I \\
\Big\updownarrow K_i & \alpha K_i & \Big\updownarrow \\
EI + S & \underset{\alpha K_s}{\overset{}{\rightleftharpoons}} & ESI
\end{array}$$
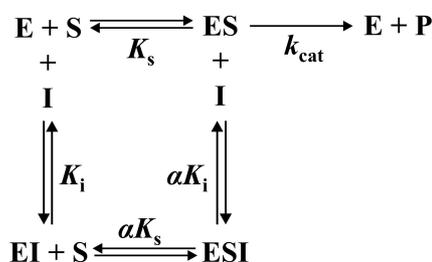
Scheme 4

In addition to the thermodynamic and kinetic constants already defined, Scheme 4 introduces the parameter α, useful to exhaustively describe the inhibition process. In particular, α quantifies the contribution to which the inhibitor affects the enzyme affinity for the substrate. If this contribution is zero, the enzyme maintains the same affinity for the substrate and $\alpha = 1$. Alternatively, if the binding of the inhibitor negatively influences further binding of substrate to the enzyme, $\alpha > 1$, becoming an infinite value for the case in which binding of substrate is completely excluded. Reversible inhibitors can be grouped in competitive, non-competitive or uncompetitive with respect to their inhibition mechanism, the latter being determined by a characteristic value of α.[14] An overview on how reversible inhibitors affect the enzyme kinetics is given in Figure 2.

A competitive inhibitor binds to the free enzyme only, competing with substrate for the binding to the enzyme active site. A reversible inhibitor cannot bind to the enzyme simultaneously to substrate so that, at any time, a fraction of the enzyme population will be bound to the substrate and will yield in ordinary catalytic activity, while another fraction will be bound to the inhibitor in a catalytically inactive form. This behaviour is described by α tending towards an infinite value. As a consequence, a higher concentration of substrate will be needed by the enzyme in order to reach half-maximal velocity. In other words, an inhibitor that conforms to a competitive inhibition mechanism will affect the apparent value of K_M that describes the enzyme-substrate pair, while V_{max} is not affected. In the steady-state condition, the initial velocity of the reaction is affected as described by the following equation:

$$v = \frac{V_{\max}[\text{S}]}{[\text{S}] + K_\text{M}\left(1 + \dfrac{[\text{I}]}{K_\text{i}}\right)} \qquad (A.14)$$

A representative plot of reaction velocity as a function of [S] in the presence of increasing concentrations of inhibitor is shown in Figure 2A. As briefly expressed in Section A.2.2, a useful tool for the analysis of the inhibition mode is the Lineweaver-Burk linearization method. In a representative Lineweaver-Burk plot the contribution of a competitive inhibitor to the kinetic properties of the system is well distinguishable (Figure 2B). In fact, the y-intercept value defining $1/V_{\max}$ is constant for all the concentrations of inhibitor tested, while the slope and the x-intercept, defining $K_\text{M}/V_{\max}$ and $-1/K_\text{M}$, respectively, are strongly affected by changing [I].

A non-competitive inhibitor can bind to both the free enzyme E and the ES complex. In this case, two equilibrium dissociation constants, rather than only one, correctly define the global affinity of the inhibitor for the enzyme. The first one, namely K_i, describes the equilibrium between I and E, while the second dissociation constant, namely αK_i and sometimes referred to as K_iu, describes the equilibrium between I and the ES complex. In non-competitive inhibition the α value can be bigger, equal or smaller than 1. In the first case, αK_i is higher than K_i, meaning that the inhibitor prefers to bind to the free enzyme, rather than the ES complex. In the case in which α is smaller than 1, αK_i is smaller than K_i, resulting in a binding preference of the inhibitor for the ES complex. The special case in which α is equal to 1 gives a αK_i value equal to K_i value. Chemically speaking, this corresponds to an equivalent affinity of the inhibitor for both the free enzyme and the ES complex. From a structural point of view, a non-competitive inhibitor negatively influences enzyme activity by interacting with a portion of the enzyme different from the active site. Hence, such an inhibitor does not compete with substrate for the active site of enzyme. As a consequence, measurements of the reaction rate in the presence of a non-competitive inhibitor would result in a strong influence on the apparent $V_{\max}$, while the apparent K_M is not affected. A general non-competitive inhibitor affects the initial velocity as described by the following equation:

$$v = \frac{V_{max}[S]}{[S]\left(1 + \dfrac{[I]}{\alpha K_i}\right) + K_M\left(1 + \dfrac{[I]}{K_i}\right)} \tag{A.15}$$

A representative plot of reaction velocity as a function of [S] measured at increasing concentrations of inhibitor is shown in Figure 2C. The resulting Lineweaver-Burk plot (Figure 2D) points out the decreasing trend of V_{max} upon increase of inhibitor concentration (the y-intercept, defining $1/V_{max}$, increases by increasing the concentration of inhibitor tested). In such an inhibition mode, the x-intercept gives information about the magnitude of α. In fact, the linear fitting would intercept above or below the x-axis in the case of $\alpha > 1$ and $\alpha < 1$, respectively.

Differently from previous cases, an uncompetitive inhibitor binds to the ES complex only. In order for the interaction between E and I to occur, formation of the ES binary complex is needed. An uncompetitive inhibitor affects the enzymatic activity determining a decrease of both apparent V_{max} and K_M values that is proportional to [I]. In uncompetitive inhibition α is always smaller than 1. This behaviour reflects a decrease of the dissociation constant that describes the equilibrium between I and the ES complex αK_i, the latter becoming the only significant term in the rate velocity equation as follows:

$$v = \frac{V_{max}[S]}{[S]\left(1 + \dfrac{[I]}{\alpha K_i}\right) + K_M} \tag{A.16}$$

The characteristic plot that describes the contribution of an uncompetitive inhibitor to the initial velocity as a function of [S] is shown in Figure 2E. The resulting Lineweaver-Burk plot (Figure 2F) yields in a series of parallel lines, where the $1/V_{max}$ value is not constant at increasing concentrations of inhibitor, while K_M/V_{max} is constant.

The global view of the different reversible inhibition modes allows one to compare the general equations that describe each mode of action (Equations A.14, A.15 and A.16). In particular, it should be noted that both competitive and uncompetitive inhibitions could

be described as special cases of non-competitive inhibition where in turn αK_i and K_i become negligible.
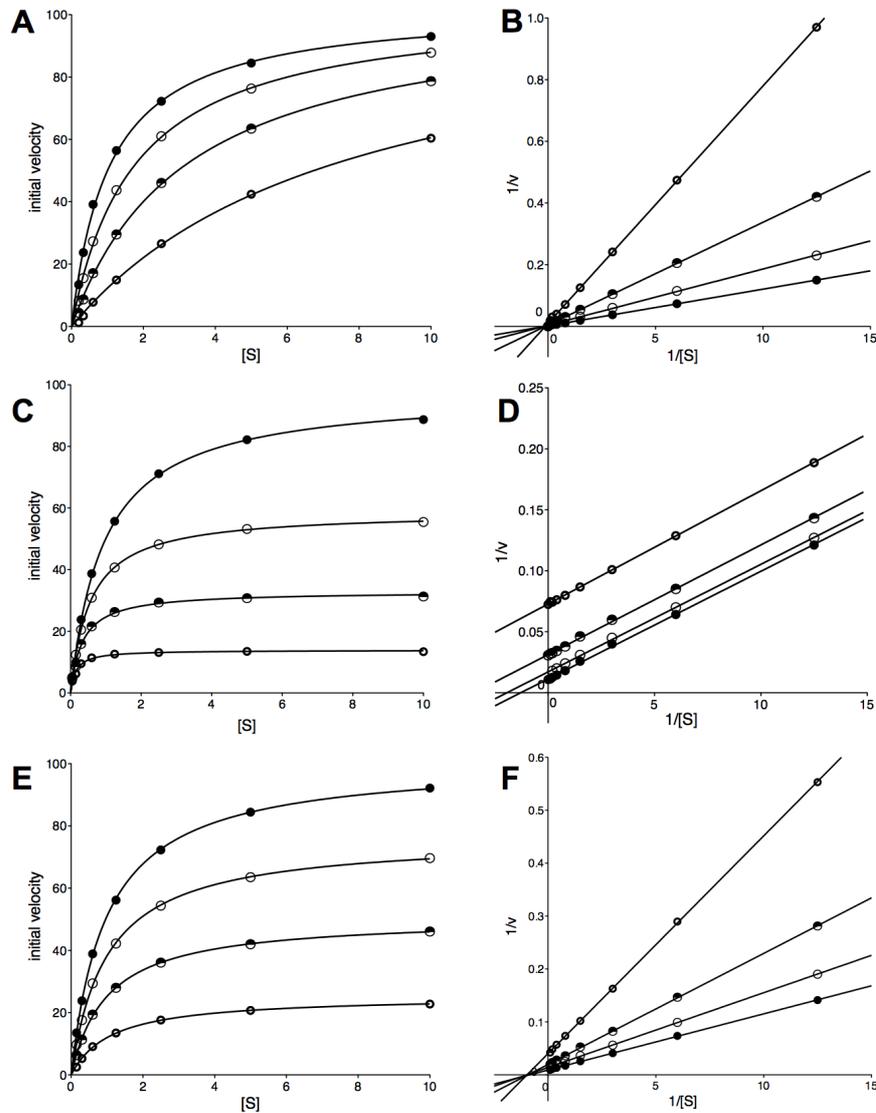
Figure 2: Steady-state velocity as a function of substrate concentration in the presence of different reversible inhibition modes. (A-B) Reaction velocity in the presence of a competitive inhibitor, untransformed and linearized data, respectively. (C-D) Reaction velocity in the presence of an uncompetitive inhibitor, untransformed and linearized data, respectively. (E-F) Reaction velocity in the presence of a non-competitive inhibitor, untransformed and linearized data, respectively. Redrawn from the book chapter: *Reversible modes of inhibitor interactions with enzymes*.[14]

A.3.2 Slow-binding inhibitors

In reversible inhibition, a binding equilibrium for an enzyme-inhibitor system is established rapidly compared with the turnover rate of the enzyme-catalysed reaction. However, a distinct class of inhibitors, called *slow-binding inhibitors*, can alter enzyme activity by means of a slow interaction with the enzyme. In slow-binding inhibition, the equilibrium between enzyme, inhibitor, and enzyme-inhibitor (EI) complex is established on a time scale of seconds to minutes, so that the inhibition strength varies with time. For this reason, slow-binding inhibitors are also referred as to *time-dependent inhibitors*. They can act by the following two mechanisms: i) simple reversible slow-binding, also referred to as mechanism A,[15] or ii) enzyme isomerization, also referred to as mechanism B,[15] which are schematically represented in Scheme 5.

Simple reversible slow-binding inhibition (Scheme 5A) involves the formation of a EI complex in a single step, whose equilibrium is regulated by smaller k_{on} and k_{off} than those regulating the ES complex formation. Hereafter, k_{on} and k_{off} will be referred to as k_3 and a k_{-3}, respectively. Thus, formation of ES complex is instantaneous, while formation of EI complex is much slower. Oppositely to the simple reversible slow binding, enzyme isomerization (Scheme 5B) involves a two-step process: first, enzyme and inhibitor form an EI complex through a rapid equilibrium governed by k_3 and k_{-3}. Next, the enzyme undergoes a slow isomerization step, governed by k_4 and k_{-4}, after which a new form of the complex, E*I, arises with increased affinity for the inhibitor. In this type of inhibition mechanism, the actual affinity has to be determined with respect to E*I complex rather than EI, so that any determination of K_i has to take into account this additional step.
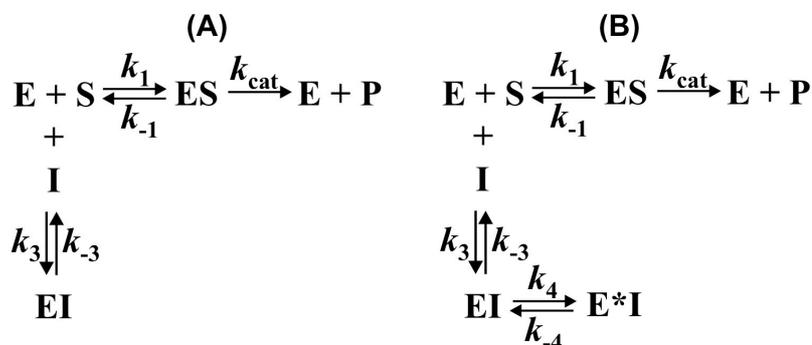
Scheme 5

Slow-binding inhibitors can be distinguished from classical reversible inhibitors by measuring the concentration of product [P] as a function of initial substrate and inhibitor concentrations over time (Figure 3).[16] A simple experimental procedure able to point out a slow-binding inhibition mechanism is a *time-course* experiment, also known as *progress-curve experiment*. In the presence of a fast equilibrium inhibition process, product formation linearly increases over time. By contrast, in the presence of a slow-binding inhibitor, product formation over time follows the behaviour represented in Figure 3A. In the initial phase of the reaction, [P] increases linearly with time. Here, the equilibrium between enzyme and inhibitor has not been established, so that the slope of the curve yields the initial velocity (v_0) of the enzyme reaction. In the intermediate phase of the time course, a significant curvature in [P] as a function of time emerges. At this stage, enzyme and inhibitor are interacting, but the system has not reached equilibrium, yet. In the late phase of the reaction, the increase of [P] with time reverts to linearity, with a slope yielding the steady-state velocity (v_s) value reached after the enzyme-inhibitor equilibrium has been established. A mathematical depiction of the progress curve as that described above is expressed by the following equation:

$$[P] = v_s t + \frac{v_0 - v_s}{k_{obs}} \left[1 - \exp\left(-k_{obs} t\right) \right] \tag{A.17}$$

Where k_{obs} is the rate constant that regulates the conversion from the initial velocity phase to the steady-state velocity phase. As stated above, slow-binding inhibition occurs in a time scale within seconds and minutes. However, in some instances the establishing equilibrium is so slow that it may be convenient to determine the residual activity of the enzyme, by means of measuring the initial velocity of the reaction under study before and after the pre-incubation of enzyme and inhibitor, as a function of pre-incubation time (Figures 3B and 3C). The resulting experimental curve would thus be described by the following exponential decay expression:

$$v_t = v_s \exp(-k_{obs} t) \tag{A.18}$$

Where v_t is the reaction velocity measured after a pre-incubation time t and v_s is the reaction velocity at pre-incubation time equal to 0.
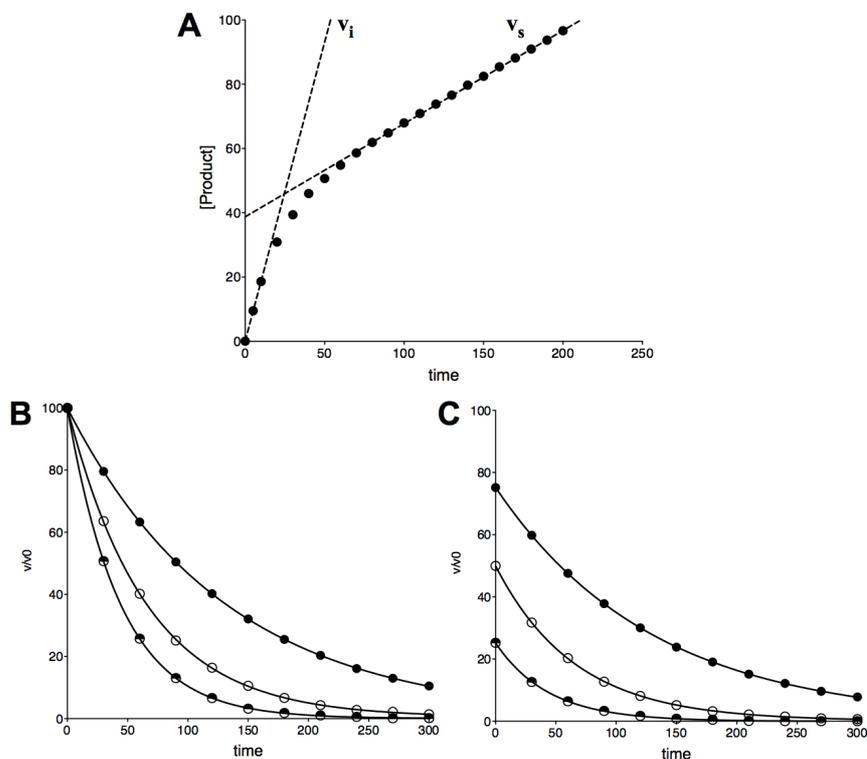
Figure 3: Kinetic profile of a slow-binding inhibition (A) Typical progress curve for an enzyme reaction in the presence of a slow binding inhibitor. (B) Typical pre-incubation experiments in the presence of a slow-binding inhibitor that conforms to a binding mechanism A. (C) Typical pre-incubation experiments in the presence of a slow-binding inhibitor that conforms to a binding mechanism B. Redrawn from the book chapter: *Slow-binding inhibitors.*[16]

As we can notice by comparing the two alternative experimental setups, the process is described by a pseudo-first order time constant k_{obs} that governs how fast the inhibition process takes place and that depends on concentration of inhibitor tested. In order to discern whether the inhibition is governed by a simple reversible or an enzyme isomerization mechanism, the dependence of k_{obs} as a function of [I] must be defined (Figure 4).

In the case of a simple reversible slow-binding inhibitor, k_{obs} linearly depends on [I] by the following equation:

$$k_{obs} = k_3[I] + k_{-3} \tag{A.19}$$

where k_3 and k_{-3} can be graphically determined from the slope and the y-intercept of the resulting plot, respectively (Figure 4A). It has to be noted that k_3, as well as the resulting K_i ($K_i = k_{-3}/k_3$) is an apparent value, since it may be influenced on the concentration of substrate.

In the case of an enzyme-isomerization mechanism, k_{obs} hyperbolically depends on [I] as follows:

$$k_{obs} = k_{-4} + \frac{k_4}{1 + \left(\dfrac{K_i^{app}}{[I]} \right)} \tag{A.20}$$

where K_i^{app} represents the apparent inhibition constant value for the EI complex (fast equilibrium). The resulting plot will give k_{-4} and $k_4 + k_{-4}$ from non-zero y-intercept and the maximum asymptotic value of k_{obs}, respectively (Figure 4B). Moreover, the concentration of inhibitor yielding a half-maximal value of k_{obs} corresponds to K_i^{app}. This value can be then correlated to the true value of K_i by determining the inhibition mode of the molecule (i.e. competitive, non-competitive or uncompetitive) and by relating K_i to K_i^*, the inhibition constant that regulates the high affinity E*I complex, by the following formula:

$$K_i^* = \frac{K_i}{1 + \left(\dfrac{k_4}{k_{-4}} \right)} \tag{A.21}$$
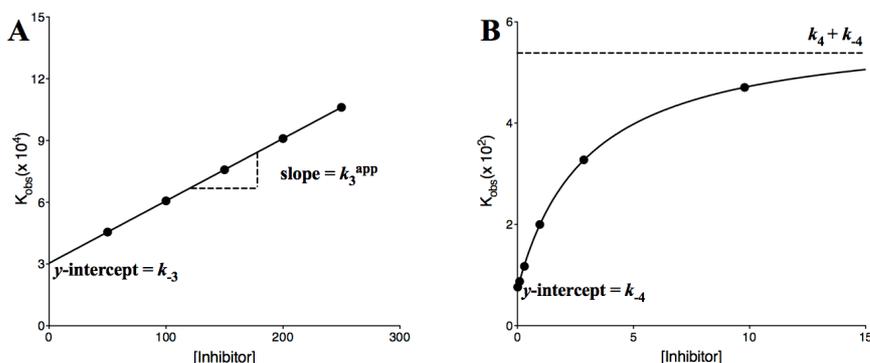
Figure 4: Plot of k_{obs} as a function of [I] in the case of a slow-binding inhibitor that conforms to mechanism A (A) and to mechanism B (B), respectively. Redrawn from the book chapter: *Slow-binding inhibitors*.

A.3.3 Irreversible inhibitors

The inhibition types discussed so far have reversibility as a peculiar common feature. Unlike the previous cases, *irreversible inhibitors*, also known as *inactivators*, are molecules that covalently bind to an enzyme, leading to its catalytic inactivation (Figure 5).[17] Kinetically, the formation of a covalent bond is a slower process as compared to the non-covalent interactions influencing equilibrium between an enzyme and a reversible inhibitor, therefore irreversible inhibition can be depicted as a particular case of slow-binding inhibition. The distinctive trait of an irreversible inhibitor is its stoichiometric inhibitory action on the enzyme so that, whenever the concentration of the inhibitor is equal or higher than the [E], the steady-state velocity reached during a progress curve experiment will be zero (Figure 5A). The equation that describes the formation of product as a function of time for an irreversible inhibitor is a simplified form of that used for the slow binding inhibition, as follows:

$$[\text{P}] = \frac{v_0}{k_{\text{obs}}} \left[1 - \exp\left(-k_{\text{obs}} t\right) \right] \tag{A.22}$$

As for the reversible slow-binding inhibition, the inactivation rate of an irreversible inhibitor is determined by the dependence of k_{obs} on the inhibitor concentration. In particular, there are three main mechanisms through which irreversible inhibition takes place (described in Scheme 6): i) non-specific affinity labelling, ii) quiescent affinity labelling and iii) mechanism based inactivation.

(non-specific affinity label) $\text{E} + \text{I} \xrightarrow{\;k_{\text{inact}}[\text{I}]\;} \text{E-I}$

(quiescent affinity label) $\text{E} + \text{I} \underset{k_{-3}}{\overset{k_3[\text{I}]}{\rightleftharpoons}} \text{EI} \xrightarrow{\;k_4\;} \text{E-I}$

(mechanism-based) $\text{E} + \text{I} \underset{k_{-3}}{\overset{k_3[\text{I}]}{\rightleftharpoons}} \text{EI} \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} \text{EA} \xrightarrow{\;k_5\;} \text{E-A}$
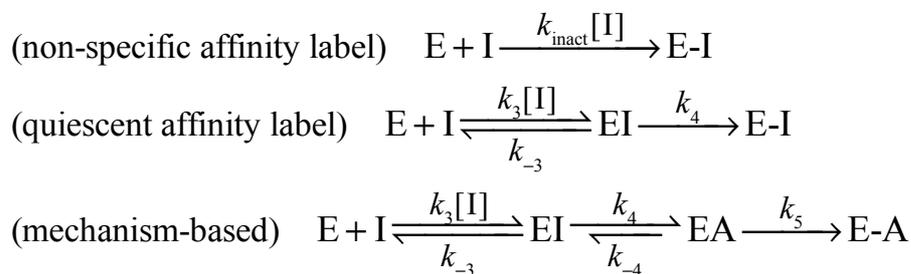
Scheme 6

Generally, a non-specific affinity label is a molecule that is intrinsically prone to be chemically modified, as it contains chemical functionalities highly reactive towards complementary functional groups belonging to the enzyme. Thus, such an inactivator covalently modifies the enzyme in a non-specific manner, by binding many amino acid groups. Inactivation occurs through a single inactivation step, with k_{obs} linearly dependent on [I], as shown in Figure 5B. The slope of the resulting linear fit gives a measure of the effectiveness of the inhibitor and it is reported in literature as k_{inact}/K_I. k_{inact} and K_I are usually complex rate constants. A full description of these parameters is available in literature[18]. Besides their complexity, k_{inact}/K_I ratio is the key metric that chemists and biologists use in order to compare the inactivation strength of a series of molecules during Structure-Activity Relationship (SAR) experiments.

Slightly different from the non-specific affinity label, a quiescent affinity label is a molecule presenting two features: i) functionalities that reversibly interact with the target enzyme, and ii) weak electrophilic functionalities with a non-intrinsic reactivity. Such a molecule first reversibly interacts with the enzyme. In this contest, the enzyme can orient the molecule so that the covalent binding reaction can occur. This kind of inactivation is more selective than the nonspecific affinity label.

Finally, a mechanism-based inactivator relies on its transformation, caused by a catalytic event performed by the enzyme, in order to become a species that may act either as an affinity label, a transition state analogue, or a tight binding reversible inhibitor.[19] Since these molecules have to be processed by the active site during the first step, they necessary act as competitive inhibitors and are also named *suicide substrates*. The need of an initial reaction with the enzyme makes these inhibitors high selective, therefore quite useful as potential drugs.

As described above, in the quiescent affinity label and in the mechanism-based mechanisms the reaction passes through a two-step inactivation process. In both cases, k_{obs} hyperbolically depends on the concentration of inhibitor (Figure 5C), resulting in a behavior described by the following equation:

$$k_{obs} = \frac{k_{inact}}{1 + \left(\dfrac{K_I}{[I]}\right)} \qquad (A.23)$$

As for the previous case, the effectiveness of inactivation is given by the ratio of k_{inact} on K_I.
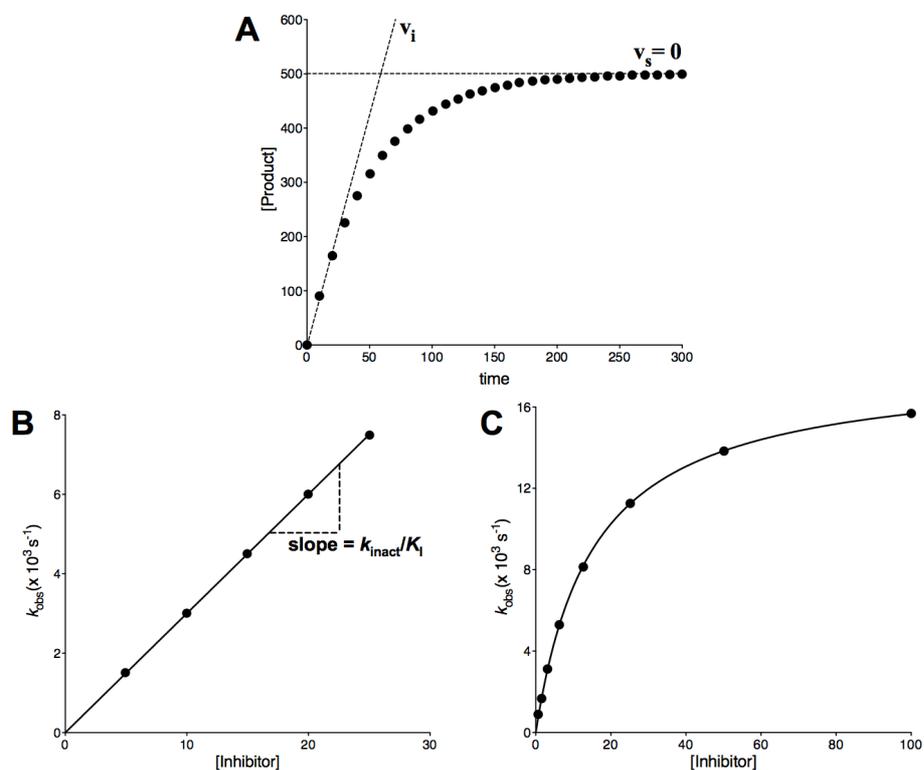
Figure 5: Inactivation kinetics of irreversible inhibitors. (A) Progress curve for an enzyme reaction in the presence of an irreversible inhibitor. (B) A plot of k_{obs} as a function of [I] for in the case of a non-specific affinity label. (C) A plot of k_{obs} as a function of [I] in the case of an irreversible inhibitor that conforms to a two-step inactivation mechanism, either quiescent affinity label or a mechanism-based. Redrawn from the book chapter: *Irreversible enzyme inactivators*.[17]

REFERENCES

1. E. Buchner (1897), *Ber. Dtsch. Chem. Ges.*, **30**:1110.
2. J. B. Sumner (1926), *J. Biol. Chem.*, **69**:435.
3. A. Wurtz (1880), *C. R. Hebd. Seances Acad. Sci.*, **91**:787.
4. A. J. Brown (1902), *J. Chem. Soc., Trans.*, **81**:373.
5. V. Henri (1903), Librairie Scientifique A. Hermann, Paris.
6. L. Michaelis and M. L. Menten (1913), *Biochem. Z.*, **49**:352.
7. G. E. Briggs and J. B. S. Haldane (1925), *Biochem. J.*, **19**:338.
8. B. Krajewska (2009), *J. Mol. Catal. B Enzym.*, **59**:9.
9. R. A. Copeland (2000), in *Enzymes: a practical introduction to structure, mechanism, and data analysis, 2nd Edition*, John Wiley & Sons, Inc., New York, USA, pp. 188.
10. R. A. Copeland (2013), in *Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists, 2nd Edition*, John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 25.
11. H. Lineweaver and D. Burk (1934), *J. Am. Chem. Soc.*, **56**:658.
12. M. A. Navia and M. A. Murcko (1992), *Curr. Opin. Struct. Biol.*, **2**:202.
13. R. Prasad and J. F. Power (1995), in *Adv. Agron.*, ed. D. Sparks, Academic Press, pp. 233.
14. R. A. Copeland (2013), in *Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists, 2nd Edition*, John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 57.
15. J. F. Morrison (1982), *Trends Biochem. Sci.*, **7**:102.
16. R. A. Copeland (2013), in *Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists, 2nd Edition*, John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 203.
17. R. A. Copeland (2013), in *Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists, 2nd Edition*, John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 345.
18. R. A. Copeland (2000), in *Enzymes: a practical introduction to structure, mechanism, and data analysis, 2nd Edition*, John Wiley & Sons, Inc., New York, USA, pp. 318.
19. L. M. Szewczuk, L. Forti, L. A. Stivala and T. M. Penning (2004), *J. Biol. Chem.*, **279**:22727.

APPENDIX B

MACROMOLECULAR CRYSTALLOGRAPHY

B.1 A brief history of macromolecular crystallography

2014 was celebrated as the International Year of Crystallography (IYCr-2014) to commemorate the centenary of X-ray crystallography discovery, dated back in 1914. In that period Max von Laue, William Henry Bragg and William Lawrence Bragg were awarded Nobel prizes for Physics for having pioneered the use of X-rays in obtaining structural information from crystals.[1, 2] Since that time, X-ray crystallography has come a long way. About biological macromolecular X-ray crystallography (usually referred to as bio-crystallography), a great number of brilliant discoveries have been achieved in a century. Among the others, three breakthroughs deserve a special mention. In 1926, Sumner demonstrated that enzymes were proteins and that they could be crystallized retaining their functions.[3] In 1958, the first three-dimensional model of a protein, myoglobin, was published.[4] In the same period, Watson and Crick structurally described for the first time the double helix of DNA.[5]

Parallel to the exponential growth of bio-crystallography, the establishment of the Protein Data Bank (PDB)[6, 7] as a single repository for structural models obtained by X-ray crystallography (and later for structures obtained by NMR spectroscopy, electron microscopy and other techniques) provided a unique resource for scientific community and determined a further increase of the structural information available on macromolecules. Nowadays, X-ray crystallography has become a technique of choice in the structural determination of proteins, as well as protein-protein/protein-nucleic acids complexes. Additionally, great developments in protein production methods (i.e. recombinant DNA techniques), establishment of crystallization facilities and

improvements in the hardware of X-ray sources (i.e. synchrotron radiation and free electron laser) have pushed X-ray crystallography towards the exploration of structure-activity relationship (SAR) between proteins and ligands, such as substrates and inhibitors. In this context, elucidation of enzyme mechanisms, as well as specificity of protein-inhibitor interactions, led X-ray crystallography to become an important tool in in structure-based drug discovery and development. In this chapter the bases of macromolecular crystallography applied to proteins will be overviewed.

B.2   Protein crystallization

X-ray protein crystallography relies on the obtainment of a protein in a crystalline form. Crystallization of proteins is a physico-chemical process that proceeds through two steps: i) nucleation, a phase transition by which the protein passes from being in solution to an ordered solid state yielding assemblies referred to as *critical nuclei*, and ii) crystal growth.[8] The crystallization process occurs in super-saturated conditions realized by reducing protein solubility, as described in Figure 1. Variation of experimental parameters as temperature, pH and ionic strength can induce super-saturation. However, this condition is frequently achieved by using precipitant agents, which reduce protein solubility by competing for water molecules, by increasing attraction between macromolecules, or by producing volume-exclusion effects.[8]
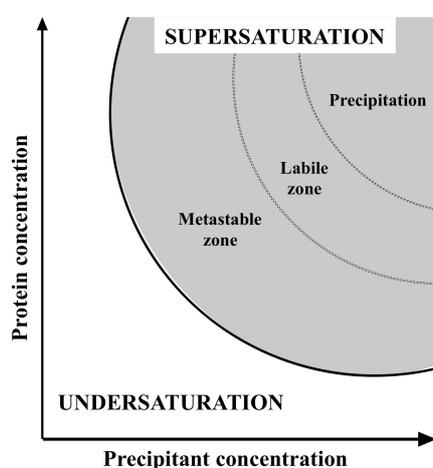
Figure 1: The phase diagram shows an undersaturated region, in which the protein is soluble, and a supersaturated region, characterized by a non-equilibrium state in which the protein is beyond its solubility limit. The supersaturated region is divided in three zones. In the precipitation zone the conditions are too extreme and amorphous aggregation occurs. In the labile area nuclei can grow, while in the metastable area nuclei will develop into crystals. Equilibrium conditions are re-established by the formation of a solid state, either amorphous precipitate or crystals, until the saturation limit is attained.

Growing protein crystals is usually a two-step approach. First, initial crystallization conditions are pursued by the use of commercial crystallization screenings that combine potential crystallization solutions and conditions. At this stage crystal dimensions and diffraction properties are usually not satisfying. Therefore, a systematic variation of the initial crystallization parameters is carried out, by the use of trial matrices, in order to optimize crystallization conditions and obtain crystals suitable for X-ray diffraction experiments.[9]

Nowadays, vapour diffusion and batch method are the most common approaches used for growing protein crystals.[8] Vapour diffusion involves the establishment of supersaturated conditions upon equilibration of the ionic strength between a drop, containing the protein solution and some amounts of a precipitant, and the reservoir, containing a higher concentration of precipitant.[10] On the other hand, the batch method is largely used for initial screening of crystallization conditions and it consists in adding the precipitant at supersaturated conditions and covering the solution with silicone oil.[11] Other crystallization methods are dialysis and liquid-liquid diffusion.[8]

Growing protein crystals in the presence of ligands may be desirable in at least two cases. First, for the initial phase estimation of structure factors, the insertion of heavy atoms is often needed. Moreover, the use of X-ray crystallography in the process of drug discovery implies the obtainment of a protein bound to drug-like molecules, in order to have structural information on their interaction. Although details on these subjects will be discussed in Section B.7 and in Chapter 1, respectively, here it is worth mentioning that these approaches can be accomplished by either co-crystallization, in which protein and ligand are pooled in the same crystallization drop in order for a co-crystal to grow, or soaking, in which a protein crystal is immersed in a ligand solution in order for the ligand to be bound into the protein crystal.[10]

B.3   Crystals and symmetry

In a protein crystal, as a particular case of a more general crystalline form, a protein (*motif*) is regularly repeated in the space by translational and symmetry operations. A basic tridimensional entity called *unit cell*, whose dimensions are defined by three vectors

a, *b* and *c*, while its angles are named α, β and γ, undergoes a discrete translation in the crystal.[12] A single unit repeated through symmetry operations forms the inner content of the unit cell. The smallest part of the unit cell not showing any symmetry element is referred to as *asymmetric unit*.[12] The only symmetry operation occurring within the unit cell of a protein crystal is the rotation around an axis (the other two possible symmetry operations, reflection across a mirror plane and inversion through a point do not occur because of protein chirality). In general, the particular set of symmetry operations that order the content of a unit cell define 32 *crystallographic point groups* or *crystal classes*.[13] However, due to protein chirality, in protein crystallography only 11 crystallographic point groups occur, the latter describing 7 *crystal systems*. Moreover, the translational repetition of the unit cell in space originates a *crystalline lattice*, where each *lattice point* is defined as a specific arrangement of atoms reproduced many times in the crystal, each one having the same environment and orientation with respect to each other point in the lattice. The unit cell and the related lattice points can undergo a translation along the tridimensional lattice (*x*, *y*, *z*) equal to the length of each cell vector, *a*, *b* and *c*. Depending on the crystal system and how translations occur, 14 *Bravais lattices* can be produced. Location of the lattice points in the unit cell upon its translation in the lattice defines four centring types for the unit cell: i) primitive (P), ii) face-centred (F), iii) base-centred (C), and iv) body-centred (I).[13] Additionally to the rotation symmetry, the screw axes are an internal symmetry operation in protein crystals, where a rotation is combined with a translation parallel to the rotation axis. The overall combination of the point groups with the Bravais lattices and the internal symmetry operations originate 230 *space groups*.[13] In protein crystallography this number is reduced to 65 chiral space groups.[12]

LATTICE PARAMETERS	CRYSTAL SYSTEM	POINT GROUP (HERMANN – MAGUIN NOTATION)	CELL TYPE (BRAVAIS LATTICE)	SPACE GROUP
$a \neq b \neq c$ $\alpha \neq \beta \neq \gamma \neq 90°$	TRICLINIC	1	P	P1
$a \neq b \neq c$ $\alpha = \gamma = 90°$ $\beta \neq 90°$	MONOCLINIC	2	P	P2, P2$_1$
			C	C2
$a \neq b \neq c$ $\alpha = \beta = \gamma = 90°$	ORTHORHOMBIC	222	P	P222, P222$_1$, P2$_1$2$_1$2, P2$_1$2$_1$2$_1$
			I	I222, I2$_1$2$_1$2$_1$
			C	C222$_1$, C222
			F	F222
$a = b \neq c$ $\alpha = \beta = \gamma = 90°$	TETRAGONAL	4	P	P4, P4$_1$, P4$_2$, P4$_3$
			I	I4, I4$_1$
		422	P	P422, P42$_1$2, P4$_1$22, P4$_1$2$_1$2, P4$_2$22, P4$_2$2$_1$2, P4$_3$22, P4$_3$2$_1$2
			I	I422, I4$_1$22
$a = b \neq c$ $\alpha = \beta = 90°$ $\gamma = 120°$	TRIGONAL	3	P	P3, P3$_1$, P3$_2$
			R*	R3
		32	P	P312, P321, P3$_1$12, P3$_1$21, P3$_2$12, P3$_2$21
			R*	R32
	HEXAGONAL	6	P	P6, P6$_1$, P6$_5$, P6$_2$, P6$_4$, P6$_3$
		622	P	P622, P6$_1$22, P6$_5$22, P6$_2$22, P6$_4$22, P6$_3$22
$a = b = c$ $\alpha = \beta = \gamma = 90°$	CUBIC	23	P	P23, P2$_1$3
			I	I23, I2$_1$3
			F	F23
		432	P	P432, P4$_2$32, P4$_3$32, P4$_1$32
			I	I432, I4$_1$32
			F	F432, F4$_1$32

* R=Rhombohedrally-centered: lattice points on the cell corners with two additional points along the longest body diagonal (only applies for the hexagonal crystal family)

B.4   How to produce X-rays for protein crystallography

Three devices can be used for the production of X-rays for protein crystallography: X-ray tubes, rotating anodes and synchrotron radiation. X-rays produced by either an evacuated tube or rotating anodes are usually used as laboratory sources. X-ray tubes consist of a cathode from which electrons are emitted and accelerated across the vacuum towards a metal anode. Impact of the electrons on the anode causes a conversion of kinetic energy into X-rays, producing: a) a low intensity continuous spectrum of X-rays (the so-called bremsstrahlung spectrum or *white radiation*), and b) characteristic high intense emission lines that depend on the anode material. The characteristic wavelength of the X-rays emission is then selected by a monochromator and the resulting monochromatic X-rays are collimated and focused onto the crystals. A rotating anode exploits a similar procedure, but the anode target rotates at high speed to efficiently dissipate heat, so that a more intense X-ray beam can be produced. The wavelength produced by both X-ray tubes and rotating anodes is fixed by the choice of anode target material and not tunable. The most common element used as an anode is copper, with a $\lambda=1.54$ Å.

Synchrotron radiation produces more intense X-rays, whose wavelength is tunable, a useful feature to minimize absorption or perform multiple anomalous diffraction (MAD) experiments (details of the technique are described below). In order for X-rays to be produced, electrons (or positrons) are first accelerated in a linear accelerator and then injected and maintained as bunches in a storage ring. Magnetic insertion devices (IDs), such as bending magnets, wigglers and ondulators, deviate the path of the particles while they travel around the ring. Change in particles trajectory causes the emission of a narrow fan of polychromatic radiation over a tightly defined angle tangential to the ring. Therefore, radiation is tunable by selecting the appropriate wavelengths with monochromators. A new-generation source of X-ray synchrotron radiation is the free electron laser (FEL). In a FEL, a beam of electrons linearly accelerated to near-light speed passes through a series of IDs, being forced to wiggle transversely along a sinusoidal path within the ID. Change in electrons trajectory results in the release of incoherent monochromatic photons. Bunching of electrons occurs through a self-organization of electrons in relativistic conditions known as collective instability,[14] so that the radiation emitted by bunched electrons is coherent and reaches a power several

orders of magnitude higher than that of the circular storage rings. The wavelength of the radiation emitted is tunable by changing the energy of the electron beam or the magnetic-field strength of the insertion devices.

B.5 Principles of crystallography

B.5.1 Principles of X-ray diffraction

X-ray diffraction by crystals can be treated as a special case of reflection by using the Bragg's model, as shown in Figure 2. The incident beam is reflected from sets of equivalent and parallel planes in the crystal, which are defined by the *lattice indices h, k* and *l*, each one representing the number of planes, describing that particular *hkl* set, that intersect the unit cell along *a*, *b* and *c*, respectively. X-ray diffraction can be explained, using the description above, by the Bragg's law, as follows:

$$2d_{hkl}\sin\theta = n\lambda \qquad\qquad (B.1)$$

In this equation, d_{hkl} is the interplanar distance of a set of *hkl* planes and θ is the angle of the reflected beam with respect to the *hkl* set of planes. According to this model, diffraction occurs when the reflection θ angle is such that the difference in path length for X-rays reflected from two successive planes of the same *hkl* set is equal to an integer number of wavelengths (being the wavelength assumed as the one of the incident beam). In that case, X-rays are reflected in phase with each other, constructively interfering and giving rise to a diffracted beam. For other reflected θ angles, waves reflected by successive planes will destructively interfere, thus no diffraction beam will be produced.
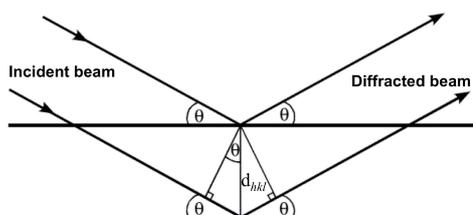
Figure 2: Diffraction of X-rays and description of the Bragg's model in real space.

Bragg's Law as defined above is an explanation of how diffraction arises in real space. However, a good point is to look at diffraction in the *reciprocal space*. In the reciprocal space a crystal can be defined in terms of a *reciprocal lattice*, as it has been already defined in terms of crystalline lattice in real space. A geometrical reconstruction of the Bragg's model in the reciprocal space is the Ewald's sphere (Figure 3). This is a useful tool because each reflection arising from the interaction between the incident X-rays and the planes in the crystal gives a diffraction pattern that is reciprocal to the direct lattice. In the reciprocal space, dimensions of the unit cells are $a^*=1/a$, $b^*=1/b$ and $c^*=1/c$, so that an *ab* plane of a unit cell in real space will yield in an a^*b^* plane in the reciprocal space. A reciprocal lattice can be constructed starting from any lattice point O, the latter being a lattice point also in the real space. When a X-ray beam (in the XO direction) impacts on the crystal along the a^*b^* plane, a circle of radius $1/\lambda$ (with λ being the incident X-ray wavelength) can be drawn by centring it in the point C on the OX direction and passing through O. This circle represents the wavelength of the incident X-ray beam in the reciprocal space. In such a situation, the reciprocal lattice point P lies on the circle. The angle PBO corresponds to the reflection angle θ and the triangle PBO, being inscribed in a semicircle, it is a right triangle, so that:

$$\sin\theta = \frac{OP}{BP} = \frac{OP}{2/\lambda}$$

(B.2)

Since P is a reciprocal lattice point, the length of OP is $1/d_{hkl}$ and the expression above can be rearranged as follows:

$$2d_{hkl}\sin\theta = \lambda$$

(B.3)

That is the Bragg's law with n=1. If we now rotate the reciprocal lattice about O (by rotating the crystal), point P' will cross the circle. In this new situation, P' has *h'k'l'* indices points, so that:

$$2d_{h'k'l'}\sin\theta = \lambda \qquad\qquad (B.4)$$

Concluding, the Ewald's sphere reconstruction provides an understanding for whether reflections will occur or not. Whenever a crystal is rotated so that a reciprocal lattice point intersects the Ewald's sphere, Bragg's law is satisfied and diffraction will take place. Because a diffraction data collection is usually carried out by rotating a crystal while it is irradiated by X-rays, many reciprocal lattice points will cross the Ewald's sphere, each one giving rise to a reflection.[15] The reflection originated by the reciprocal lattice point P_{hkl} will give the *hkl* reflection that is, according to the Bragg's model, the reflection of a *hkl* set of planes in real space. Bragg's model assumes that the number and direction of the reflections are the result of the dimensions of the unit cell, whereas the intensity I_{hkl} of each reflection is the result of the content of the unit cell, in particular of the electron density belonging to a set of equivalent *hkl* planes.
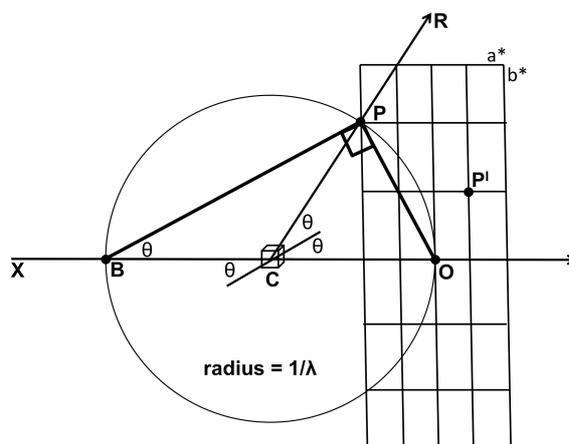
Figure 3: Diffraction of X-rays in the reciprocal space described by the Ewald's sphere.

B.5.2 The structure factor

The intensity I_{hkl} of a reflection is, numerically, the squared amplitude of the *structure factor* F_{hkl} ($|F_{hkl}|^2$), with the latter being a periodic function of the electron density distribution in the unit cell, as follows:

$$F_{hkl} = \sum_{j=1}^{n} f_j \exp\left[2\pi i\left(hx_j + ky_j + lz_j\right)\right] \qquad (B.5)$$

The term f_j is called the *atomic scattering factor* of an atom j. The atomic scattering factor is a mathematical description of an atom as a centrosymmetric scatterer, and it is related to the number and position of the electrons in its orbitals. The kinetic energy available at a certain temperature different rom 0 K causes the periodic change in the position of atoms (vibration) around an average position. This vibration, whose amplitude is temperature-dependent, is called *B-factor* or temperature factor and is described by the following expression:

$$B_j = 8\pi^2 U_j^2 \qquad (B.6)$$

where U_j^2 is defined as the mean-square displacement of atom j from its rest position and is expressed in $Å^2$. The structure factor F_{hkl} describes a reflection *hkl* given by the contribution of the atoms located in the unit cell and assumed as spherical clouds of electron density. The contribution of each atom j to F_{hkl} depends on two factors: i) the atom type contributes to the amplitude of F_{hkl}, while ii) the position (x_j, y_j, z_j) of the atom j in the unit cell, contributes to the phase of F_{hkl}, being the latter a periodic oscillation function. The total structure factor F_{hkl} can be alternatively described considering the electron density (ρ) belonging to each atom in the unit cell as a function of the unit cell volume, thus resulting in F_{hkl} as the total sum of contribution of each volume element of the unit cell to the total scattering, as follows:

$$F_{hkl} = \int_h \int_k \int_l \rho(x,y,z)\exp\left[2\pi i\left(hx + ky + lz\right)\right] dxdydz \qquad (B.7)$$

The electron density is the Fourier transform of the structure factor, hence:

$$\rho(x,y,z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} \exp\left[-2\pi i\left(hx + ky + lz \right) \right] \qquad \text{(B.8)}$$

In this equation, V is the unit cell volume. This expression indicates how to obtain the electron density from the measurement of all the reflections resulting from an X-ray diffraction experiment. However, since structure factors describe the diffraction of an electromagnetic wave, three parameters are needed for a complete description: amplitude, frequency and phase. The amplitude of structure factors has already been discussed and is a measurable amount. Frequency corresponds to that of the incident X-ray beam, since X-ray diffraction entails elastic scattering (electrons sensing the oscillating electric field of the incident X-ray beam will oscillate with the same frequency). The phase of the structure factors, that is described in Equations B.7 and B.8 by the coordinates x, y and z, is the main crystallographic problem to be solved. In fact, by performing a diffraction experiment phases of all the structure factors are lost, because the x, y, and z position of each atom (and thus electrons) within the crystal and with respect to the source of the X-ray beam is an unknown in the experimental set up. In Section B.6 an overview of the experiments that are usually performed in order to estimate the structure factor phase and reconstruct the electron density of a protein structure is briefly given.

B.6   Recap of data collection, processing and reduction

Nowadays, an X-ray diffraction experiment is performed by collecting diffraction images upon crystal rotation in order to fulfil Bragg's law crossing as many reciprocal lattice points as possible.[15] Diffraction images are usually recorded by new generation semiconductor charge-coupled detectors (CCD) or active pixel sensors in complementary metal–oxide–semiconductor detectors (CMOS). In order to minimize radiation damage, due to the absorption of X-rays by the protein and the solvent molecules with the resulting loss of electrons and formation of free radicals, data collection is performed at cryogenic temperature (100 K) by using a cryo-cooling nitrogen stream flowing onto the crystal during the experiment.[16] Once data collection is completed, the recorded diffraction images undergo a data processing and reduction protocol through some

sequential steps. The essence of this process is to transform and relate each two dimensional section of the reflection, recorded on the surface of the detector in multiple images, to a single value for each h, k, l value. Assuming that all data collection parameters (crystal to detector distance, crystal rotation angle, etc.) are available, the first step is the determination of the crystal lattice and cell parameters. Next, the intensity (I_{hkl}) of each reflection is measured with its standard deviation σ_I, and all symmetry-related observations of a given reflection are grouped together in unique reflections. The data processing and reduction processes further yield a statistical treatment describing the overall data quality. In general, final statistical parameters are given with respect to either the number of the recorded images or the resolution of the data. The agreement between the intensities of symmetry-related reflections are quantified by R-factors such as R_{merge} and R_{meas}.[17, 18] In particular, R_{merge} measures the spread of n independent measurements of the intensity of the same reflection (I_{hkl}) around their average value, while R_{meas} is an adjusted version of R_{merge} that takes into account multiplicity, another important parameter that defines the average number of measurements for that unique reflection. Other important statistical parameters obtained are data completeness and a resolution value beyond which the ratio between reflection intensities and standard errors is smaller than 2.0. Altogether, these parameters allow the crystallographer to judge the quality of collected data, in order to choose the correct resolution cut-off and the final number of images to use for the next step of structure refinement and model building.[19] Current software available for data processing are XDS[20] and MOSFLM,[21] while data reduction tools are usually implemented within suites for macromolecular crystallography, such as CCP4[22] and PHENIX.[23]

B.7   Initial phase estimation

There are three main diffraction experiments that can be performed in order for an initial phase value to be estimated: i) isomorphous replacement, ii) anomalous diffraction, and iii) molecular replacement.

B.7.1 Isomorphous replacement

This technique relies on the strong diffraction caused by elements with high atomic numbers (i.e. heavy atoms). The experiment consists in carrying out data collections on native crystals and isomorphous derivative crystals, in which a heavy metal has been inserted by either co-crystallization or soaking. In a typical multiple isomorphous replacement (MIR) three datasets, one for the native protein and two for the derivative crystals (each derivative crystal containing a different heavy atom), are collected. If we define, for each *hkl* reflection, F_P as the structure factor for the native protein and F_{PH}' as the structure factor for the first derivative, $F_{PH}' = F_P + F_H'$, where F_H' is the structure factor of the first heavy atom. First, determination of heavy atom position in the unit cell is carried out by a difference Patterson function[24] calculated using the intensities of each reflection for the native ($|F_P|$) and the derivative ($|F_{PH}'|$), respectively. This results in a Patterson map showing the heavy atom location in the unit cell that, in turn, allows knowing the phase for F_H'. F_P, F_{PH}' and F_H' can be described as vectors and their relationship can be shown by the Argand diagram in Figure 4, where the structure factor for the native protein is described as the following vector difference:

$$F_P = F_{PH}' - F_H'$$
(B.9)

The amplitude of the three vectors (their lengths in the plane) is known, as well as the phase of F_H'. By drawing a vector with module equal to $-F_H'$ at the origin of the plane, the circle of radius F_{HP} centered on the head of the $-F_H'$ vector defines all the points which give the vector difference $|F_{HP}'|-F_H'$. Next, another circle of radius F_P centered on the origin can be drawn in order for it to intersect the $|F_{HP}'|-F_H'$ circle in two points. The resulting two F_P^a and F_P^b vectors, drawn from the origin to the intersection of the two circles, represent the two possible phase solutions for the structure factor *hkl* (Figure 4A). In order to unambiguously resolve this phase uncertainty, an analogous experiment with a second derivative crystal is necessary. In fact, a similar trigonometric approach (this time using F_H'' and FP_H'') would give two new phase values for the structure factor of the native protein, one of them corresponding to one of the two previously determined, thus representing the correct solution (Figure 4B).
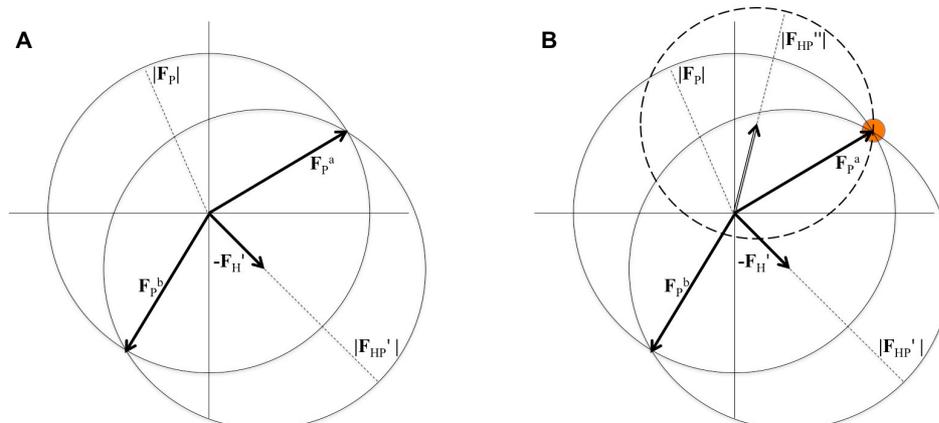
Figure 4: Argand diagram used for the estimation of initial phase by multiple isomorphous replacement. (A) Two possible solutions for the phase of F_P are available by calculating the vector solution of Equation (B.9). (B) A second derivative factor structure $|F_{HP}''|$ is represented by the points on the dashed-line circle. The correct phase value for F_P is given by the intersection of the $|F_{HP}''|$ circle with the two circles previously drawn.

B.7.2 Anomalous diffraction

Anomalous diffraction exploits heavy atoms or modified amino acids (i.e. selenomethionine) as *anomalous scatterers* in order to estimate the initial phase of a structure factor and relies on the fact that these elements have characteristic *absorption edges* in the X-ray region, so that a data collection performed at the appropriate wavelength can maximize their anomalous scattering behavior. The result of this absorption reflects on the structure factors. In fact, in the absence of anomalous scattering, structure factor vectors F_{hkl} and $F_{-(hkl)}$ have the same magnitude, so that the phase angle α is $\alpha_{hkl} = -\alpha_{-(hkl)}$ and $I_{hkl} = I_{-(hkl)}$ (a property described by the Friedel's law). In the presence of anomalous scattering, Friedel's law is no longer valid and the structure factors F_{hkl} and $F_{-(hkl)}$ differ because of a variation of scattering factors f, the latter being the result of the absorption of part of the incident X-ray beam. The scattering factor $f(\lambda)$ of a generic atom at a particular wavelength λ it is defined by the following expression:

$$f(\lambda) = f°(\theta) + f'(\lambda) + if''(\lambda) \qquad \text{(B.10)}$$

and it results from a combination of two contributions: i) the normal scattering factor (f °), which is independent of the wavelength, whereas it depends on the diffraction angle (θ) (it decreases as the scattering angle increases), and ii) the complex anomalous component having real (f ') and imaginary (f ") parts, which are dependent on the wavelength, whereas they do not depend on the scattering angle (because the anomalous scattering originates from an interaction of the incident beam with inner core electrons).[25] The scattering contribution of a structure factor can be separated in two components, a normal and an anomalous scattering,[26] so that the intensity I(h) of a reflection h (where h stands for *hkl*) measured at a particular wavelength λ is given by:

$$
\begin{aligned}
|\,^{\lambda}\mathrm{F_T}(h)\,|^2 = &\,|\,^{\circ}\mathrm{F}(h)\,|^2 + a(\lambda)\,|\,^{\circ}\mathrm{F_A}(h)\,|^2 \\
&+ b(\lambda)\,|\,^{\circ}\mathrm{F}(h)\,\|\,^{\circ}\mathrm{F_A}(h)\,|\cos\!\left[\,^{\circ}\phi(h) - \,^{\circ}\phi_{\mathrm{A}}(h)\right] \\
&+ c(\lambda)\,|\,^{\circ}\mathrm{F}(h)\,\|\,^{\circ}\mathrm{F_A}(h)\,|\sin\!\left[\,^{\circ}\phi(h) - \,^{\circ}\phi_{\mathrm{A}}(h)\right]
\end{aligned}
\tag{B.11}
$$

Where A is the anomalous scatterer and:

$a(\lambda) = (\mathrm{f'}^2 + \mathrm{f''}^2)/\mathrm{f}^{\circ 2}$

$b(\lambda) = 2(\mathrm{f'} + \mathrm{f}^{\circ})$

$c(\lambda) = 2(\mathrm{f''}/\mathrm{f}^{\circ})$

Normal scattering factors f ° are available from quantum mechanical calculations, while the anomalous scattering factors f ' and f " can be determined from a combination of theory and X-ray absorption spectra,[27] through which the wavelength-dependent factors a(λ), b(λ) and c(λ) can be determined, while the quantities |°F(h)|, |°F$_A$(h)| and °Φ(h)-°Φ$_A$(h) can be determined by trigonometric analysis. A combination of Patterson and probability approaches finally gives the solution of the phase problem as a probability distribution.[28]

The experimental procedure for phase estimation by exploiting anomalous diffraction is usually carried out through the so-called multiple-wavelength anomalous diffraction (MAD) experiments, in which two datasets are collected at two different wavelengths, the latters providing a maximal difference between f ' and f ". In this context, the tunability of

synchrotron beam lines is needed. A third wavelength is usually used to collect a dataset at lower wavelength in order to have high-resolution data to be used once the phase will be available. An extreme simplification of the experimental setup used for anomalous diffraction is the single wavelength anomalous diffraction (SAD), in which a single dataset is recorded at a wavelength with a good f" followed by density modification to solve the phase ambiguity.[29, 30]

B.7.3 Molecular replacement

Differently from the other methods, in which the estimation of initial phases is strictly experimental, molecular replacement (MR) is a trial-and-error method that enables the solution of the phase problem of an unknown protein crystal (target) by exploiting a known protein structure as an initial model. The main requirement for MR to be successful is the structural similarity between the model and the target in terms of low root mean square deviation (r.m.s.d.) of atomic position and high sequence identity.[31] The use of molecular replacement as method of choice for the structural determination of macromolecules has increased parallel to the expansion of the PDB. Nowadays up to 70% of deposited structures are determined by MR.[31, 32] Besides the possibility to follow different approaches (see below for details), MR procedure essentially consists in orienting and positioning the model in the unknown crystal in order for the resulting calculated diffraction pattern to best match the diffraction pattern collected on the target protein.

Rotation and translation operations can be carried out by following two procedures: i) a Patterson methods, or ii) a Maximum-likelihood methods. Molecular replacement has been based on the properties of the Patterson function.[24] and the resulting Patterson map. In the Patterson method the correct orientation of the model is determined by a rotation function (R), defined as a convolution between the Patterson function calculated with the squared structure factor amplitudes belonging to the target protein (P_{obs}) and the Patterson function calculated from the model rotated around an origin (P_{calc}):[33]

$$\mathrm{R(C)} = \int_U P_{\mathrm{obs}}(u) P_{\mathrm{calc}}(Cu)\,\mathrm{d}u \qquad\qquad (\mathrm{B.12})$$

where C is the operator element that describes the rotation of the P_{calc} coordinate system with respect to P_{obs}, while U is the integration volume. Rotation results are scored depending on how well the two Patterson functions superimpose their maxima. Best results from the orientation step are then used as an input for the determination of the correct position of the model in the target crystal. This step is carried out by a translation function (T),[34] defined as a convolution between the Patterson function of the observed data (P_{obs}) at the point u and the Patterson function of the model (P_{calc}) translated at the new point $u+t$.

$$T(t) = \int_{U} P_{obs}(u)P_{calc}(u+t)\,du \qquad (B.13)$$

Common macromolecular crystallography tools that exploit the Patterson method for structural determination through molecular replacement are *AMoRe*[35], *MOLREP*[36] and *X-PLOR*.[37]

Unlike the traditional Patterson methods, in maximum-likelihood methods the probability of any orientation-position combination of the model in yielding the observed structure amplitudes of the target is calculated.[31, 38] Here, the translation function is defined as a sequential positioning of any starting orientation of the model at grid points inside the unit cell. At each testing position the translation is known, so that the phase is known and the complete structure factors of the model (F_{calc}) can be calculated and subtracted by the phase contribution in order to have an unphased calculated structure factor comparable with the observed unphased structure factors (F_{obs}). The probability of observing a particular F_{obs} is given by $P(|F_{obs}|||F_{calc}|)$, with the whole procedure regulated by the so-called Rice function.[39, 40] Rice function outputs probabilities for each tested translation, selecting the best solutions. Conceptually similar to the former, maximum-likelihood rotation function assumes that the model is sequentially rotated on a grid in the angular space and the orientation that predicts the experimental data whit highest probability is selected by applying a Rice distribution. The results from both the translational and the rotational functions are then combined to give the overall probability function over the entire dataset. A macromolecular crystallography tool that exploits maximum-likelihood methods is *Phaser*.[41]

B.8   Refinement and model building

As a final step in the process of structure determination, refinement and model building are carried out in order to construct a macromolecular model that adequately explains the observed experimental data, while making physical and biochemical sense. By doing this, model parameters belonging to the initial phase estimation (obtained through MIR or MAD experiments) or the initial phasing model (used for the MR experiment) are iteratively refined against the experimental data collected during the diffraction experiment, in order for the final refined model to satisfy the best approximation between the calculated structure factors (F_{calc}) and the observed structure factors (F_{obs}). The parameters of the model to be refined are, in general, the atomic position in the unit cell, B-factors values and, occupancy (a measure of the fraction of molecules in the crystal in which an atom j actually occupies the position specified in the model). Differently from small molecules, crystallographic models of macromolecules are complex and the number of observations is not sufficient to refine the huge number of model parameters. For this reason refinement is usually carried out under the control of stereochemical and geometrical restraints that are used as prior knowledge in order to decrease the number of model parameters needed. There are two approaches able to provide useful information by performing a macromolecular refinement, the least squares and the maximum likelihood.

Least-squares is the simplest and oldest statistical method used in macromolecular refinement. It has been introduced by Konnert[42] and consists in a minimization of the differences between the observed structure factors (F_{obs}) and the calculated structure factors (F_{calc}), as follows:

$$f(p) = \sum_{hkl} \frac{(|F_{obs}| - |F_{calc}|)^2_{hkl}}{(\sigma_{obs})^2_{hkl}}$$

(B.14)

By changing the parameters of the model, a set of calculated observations that minimize the residual function is determined. Limitations of the least-squares approach, due to the need of a very complete model in order for parameters to reproduce all the observations or the assumption the measurements of the observations are the only source of errors, risk

to get trapped in local minima.[43] To overcome these limitations, maximum likelihood methods have been implemented in macromolecular refinement procedure.[44] Maximum likelihood method is a general statistical scheme[45] in which experimental observations, as well as model parameters, have an error and an uncertainty. First, for every combination of model parameters, the probability of each set by considering the prior known information belonging to the experimental data (stereochemistry, B-factors, etc.) is calculated. This probability is called *prior distribution*. Next, for every combination of model parameters, the probability that the experiment would have resulted as it did is calculated. This probability is called *likelihood distribution*. Only the sets of parameters that have high probability in both the distributions are then further combined in order to have the so-called *posterior distribution*, which includes all of the information about the values of the model parameters. The basic maximum-likelihood residual f(p) is the following:

$$f(p) = \sum_{hkl} \frac{(|F_{obs}| - |F_{calc}|)^2_{hkl}}{(\sigma_{obs})^2_{hkl} + (\sigma_{calc})^2_{hkl}} \qquad (B.15)$$

Even though the treatment is very similar with respect to the least-squares, here $|F_{calc}|$ is not the equivalent quantity calculated starting from the model parameters, but the expectation value of this quantity calculated over all the acceptable models similar to p. σ_{calc} is the width of distribution of values for $|F_{calc}|$ over all the acceptable values for p.

The progresses within and during the iterative refinement procedure can be followed by monitoring two statistical factors, *R-factor* and R_{free}, defined as a percentage:

$$R = \frac{\sum_{hkl} ||F_{obs}| - |F_{calc}||}{\sum_{hkl} |F_{obs}|} \qquad (B.16)$$

R-factor is a measure of the agreement between the refining model and the experimental data, describing how well the refined structure predicts the observed data.[46] R_{free}, is computed on a small, random sample of data never included in the refinement process as a cross validation tool in order to avoid over-fitting of the model.[47, 48] R_{free} is always

slightly greater than R-factor, but the two statistics should be comparable because a correct model should predict the experimental data with uniform accuracy. An R_{free} significantly higher than R-factor indicates an over-parameterisation of the model.

Software for final refinement, such as REFMAC[44] and *phenix.refine*,[49] as well as graphics programs used to visualize the protein model, the electron density maps and to perform manual building are included in CCP4[22] and PHENIX suites,[23] respectively. A very useful tool in the final steps of refinement is PDB_REDO,[50] a web server through which a refined structure can be automatically optimized for the following validation and deposition processes.

B.9   Structure validation and deposition

After a crystallographic structure has been refined, it has to be validated and deposited in the Protein Data Bank.[6] The access to PDB provides structural information as atomic coordinates of the deposited structures, as well as related structure factors. Availability of such structural data is important not only for scientists involved in the determination of the crystal structure of an unknown protein by using a starting model for molecular replacement, but also for biochemists interested in structure-activity relationship (SAR) of proteins or in the process of structure-based drug design (SBDD), as described in Chapter 1.

The validation process, implemented in the web server of worldwide Protein Data Bank (wwPDB, http://www.wwpdb.org/) is a needed step prior to deposition and it is important to both assess the quality of the final refined structure and allow users to discriminate between good and bad data, quantitatively expressing the reliability of the deposited model. The wwPDB validation process judges the quality of the structure. First, it analyses the reliability of the atomic model, without taking into account the diffraction data, in terms of Ramachandran torsion angles, rotameric states of amino acid side chains and close contacts between atoms. The geometry of small-molecule ligands is also assessed by comparing their structure with related small-molecule crystal structures available in the Cambridge Structural Database (CSD) (ccdc.cam.ac.uk).[51] For each evaluation criterion, a list of outliers is produced and an overall score is computed. As a second procedure it analyses the experimental diffraction data, in terms of structure

factors, without references to the atomic model, by using *phenix.xtriage*.[23] As a final step, an assessment of the fit between the atomic model and the experimental diffraction data is carried out by a number of programs.[52] As an output, a preliminary wwPDB validation report is provided to help the identification of possible issues with the atomic coordinates and/or experimental data. Once the depositor approves the preliminary wwPDB validation report, a PDB code is provided, and the submission passes to an annotation step for further consistency/error checking and subsequent final deposition in the PDB.[52]

REFERENCES

1. W. Friedrich, P. Knipping and M. Laue (1913), *Ann. Phys.*, **346**:971.

2. W. H. Bragg and W. L. Bragg (1913), *Proc. R. Soc. A*, **88**:428.

3. J. B. Sumner (1926), *J. Biol. Chem.*, **69**:435.

4. J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff and D. C. Phillips (1958), *Nature*, **181**:662.

5. J. D. Watson and F. H. Crick (1953), *Nature*, **171**:737.

6. F. C. Bernstein, et al. (1977), *J. Mol. Biol.*, **112**:535.

7. H. M. Berman, et al. (2000), *Nucleic Acids Res.*, **28**:235.

8. A. McPherson and J. A. Gavira (2014), *Acta Crystallogr. F*, **70**:2.

9. A. McPherson and B. Cudney (2014), *Acta Crystallogr. F*, **70**:1445.

10. G. Rhodes (2006), in *Crystallography made crystal clear, 3rd Edition*, Academic Press, Burlington, pp. 31.

11. N. E. Chayen (1997), *Structure*, **5**:1269.

12. G. Rhodes (2006), in *Crystallography made crystal clear, 3rd Edition*, Academic Press, Burlington, pp. 49.

13. T. Hahn and P. Paufler (1984), in *Crystal Research and Technology*, WILEY-VCH Verlag, pp. 1306.

14. R. Bonifacio, C. Pellegrini and L. M. Narducci (1984), *Opt. Commun.*, **50**:373.

15. U. W. Arndt and A. J. Wonacott (1977), *The Rotation method in crystallography: data collection from macromolecular crystals*, North-Holland Publishing Company.

16. E. F. Garman and T. R. Schneider (1997), *J. Appl. Crystal.*, **30**:211.

17. K. Diederichs and P. A. Karplus (1997), *Nat. Struct. Mol. Biol.*, **4**:269.

18. M. Weiss (2001), *J. Appl. Crystal.*, **34**:130.

19. P. R. Evans and G. N. Murshudov (2013), *Acta Crystallogr. D*, **69**:1204.

20. W. Kabsch (2010), *Acta Crystallogr. D*, **66**:125.

21. T. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell and A. G. Leslie (2011), *Acta Crystallogr. D*, **67**:271.

22. M. D. Winn, et al. (2011), *Acta Crystallogr. D*, **67**:235.

23. P. D. Adams, et al. (2010), *Acta Crystallogr. D*, **66**:213.

24. A. L. Patterson (1935), in *Zeitschrift für Kristallographie - Crystalline Materials*, p. 517.

25. W. A. Hendrickson and C. M. Ogata (1997), *Methods Enzymol.*, **276**:494.

26. W. A. Hendrickson, J. L. Smith and S. Sheriff (1985), *Methods Enzymol.*, **115**:41.

27. W. A. Hendrickson, J. L. Smith, R. P. Phizackerley and E. A. Merritt (1988), *Proteins*, **4**:77.

28. A. Pahler, J. L. Smith and W. A. Hendrickson (1990), *Acta Crystallogr. A*, **46**:537.

29. S. Benini, A. Gonzalez, W. R. Rypniewski, K. S. Wilson, J. J. Van Beeumen and S. Ciurli (2000), *Biochemistry*, **39**:13115.

30. Z. Dauter, M. Dauter and E. Dodson (2002), *Acta Crystallogr.D*, **58**:494.

31.   P. Evans and A. McCoy (2008), *Acta Crystallogr. D*, **64**:1.

32.   G. Scapin (2013), *Acta Crystallogr. D*, **69**:2266.

33.   M. G. Rossmann and D. M. Blow (1962), *Acta Crystallogr.*, **15**:24.

34.   R. A. Crowther and D. M. Blow (1967), *Acta Crystallogr.*, **23**:544.

35.   J. Navaza (1994), *Acta Crystallogr. A*, **50**:157.

36.   A. Vagin and A. Teplyakov (1997), *J. Appl. Crystal.*, **30**:1022.

37.   A. T. Brünger (1992), *X-PLOR Version 3.1 : a System for X-ray Crystallography and NMR*, Yale University Press.

38.   A. J. McCoy (2004), *Acta Crystallogr. D* **60**:2169.

39.   G. A. Sim (1959), *Acta Crystallogr.*, **12**:813.

40.   R. Read (1990), *Acta Crystallogr. A*, **46**:900.

41.   A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read (2007), *J. Appl. Crystal.*, **40**:658.

42.   J. Konnert (1976), *Acta Crystallogr. A*, **32**:614.

43.   D. E. Tronrud (2004), *Acta Crystallogr. D*, **60**:2156.

44.   G. N. Murshudov, A. A. Vagin and E. J. Dodson (1997), *Acta Crystallogr. D*, **53**:240.

45.   D. Sivia and J. Skilling (2006), *Data Analysis: A Bayesian Tutorial*, OUP Oxford.

46.   A. L. Morris, M. W. MacArthur, E. G. Hutchinson and J. M. Thornton (1992), *Proteins: Struct., Funct., Bioinf.*, **12**:345.

47.   A. T. Brünger (1992), *Nature*, **355**:472.

48.   A. T. Brünger (1997), *Methods Enzymol.*, **277**:366.

49.   P. V. Afonine, et al. (2012), *Acta Crystallogr. D*, **68**:352.

50.   R. P. Joosten, et al. (2009), *J. Appl. Crystallogr.*, **42**:376.

51.   I. J. Bruno, et al. (2004), *J. Chem. Inf. Comput. Sci.*, **44**:2133.

52.   J. Y. Young, et al. (2017), *Structure*, **25**:536.

ANNEX 1

Series: Metallobiology

Book: The Biological Chemistry of Nickel

Title: Chapter 5 – Urease

Pages: 60-97

Authors: L. Mazzei, F. Musiani and S. Ciurli

Year: 2017

Editors: D. Zamble, M. Rowinska-Zyrek, H. Kozlowski

Publisher: The Royal Society of Chemistry

CHAPTER 5

Urease

L. MAZZEI[a], F. MUSIANI[a] AND S. CIURLI*[a]

[a]Laboratory of Bioinorganic Chemistry, Department of Pharmacy and Biotechnology, University of Bologna, Viale Giuseppe Fanin 40, 40127 Bologna, Italy
*E-mail: stefano.ciurli@unibo.it

5.1   Introduction

Urease (urea aminohydrolase, E.C. 3.5.1.5) is an important nickel-dependent enzyme found in a large variety of organisms, including plants, algae, fungi, and several prokaryotes.[1,2] It is involved in the global nitrogen cycle, catalyzing the rapid hydrolytic decomposition of urea.[3,4] Urease catalysis causes an overall pH increase that has negative consequences both on human health and agriculture.[2,5] Indeed, urease is the main virulence factor of a large variety of human pathogens, such as *Helicobacter* (Chapter 16), *Proteus*, *Klebsiella*, *Pseudomonas*, and *Mycobacterium* spp. In the agricultural context, soil urease causes a significant decrease in soil fertilization efficiency using urea, due to ammonia volatilization and root damage caused by the increase in soil pH.

   Urease holds a significant place in twentieth-century science, having taken part in two fundamental landmarks of biochemistry. In 1926 James Sumner isolated and crystallized an enzyme from *Canavalia ensiformis* (jack bean) that, using his own words, "possesses to an extraordinary degree the ability to decompose urea into ammonium carbonate", thus demonstrating for the first time that enzymes are proteins.[6] This discovery led to Sumner being awarded the Nobel Prize in Chemistry in 1946. In 1975, Zerner, Dixon, and

co-workers demonstrated the requirement for nickel in jack bean urease (JBU) catalysis, providing the first model for the biological role of this metal as an enzyme cofactor.[7] Since then, ureases from several sources have been extensively investigated. Detailed information on their occurrence and roles in nature has been acquired. The availability of information on the genetic organization of urease operons, as well as the amino acid sequences and crystal structures of native ureases and urease–inhibitor complexes, has led to significant steps towards a complete understanding of the molecular basis of the catalytic mechanism, including the role of nickel in this system.[8–14] This chapter gives an account of the achievements made in the biological chemistry and structural biology of this peculiar enzyme in the past two decades.

5.2   Biological Significance of Ureases

Urea is the main product of catabolic processes of nitrogen-containing compounds in vertebrates; for example, human urine contains up to 0.4–0.5 M urea, corresponding to *ca.* 10 kg of urea produced per year.[15] The stability of urea in aqueous solutions is extremely high. Its non-enzymatic decomposition in water, which occurs through an elimination step to produce ammonia and cyanic acid, has a long half-life ($t_{1/2}$ *ca.* 40 years[16]) while its non-enzymatic hydrolysis, never observed experimentally,[17] has an estimated even longer half-life (*ca.* 520 years).[18] Together, these two aspects would cause grave environmental problems in the absence of an efficient enzymatic system able to decompose urea. The enzymatic hydrolysis of urea by urease generates two ammonium ions and a molecule of bicarbonate,[3,4] causing an overall pH increase due to product alkalinity. The enzymatic reaction has a half-time of a few microseconds, with a k_{cat}/K_M that is 3×10^{15} times higher than the rate of the uncatalyzed reaction, making urease the most efficient hydrolase known, an efficiency ascribed to the presence of two Ni(II) ions in the active site of the enzyme.[18]

Urease is widespread in nature.[1,2,19] Its general biological function is to provide organisms with a nitrogen source for growth.[12] In plants, urease is involved in metabolic pathways of nitrogen-containing compounds, such as degradation of urea after its internalization, arginase-catalyzed catabolism of arginine, and degradation of purines and ureides.[19–22] Moreover, urease plays a defense function against plant pathogens with a mechanism that is unrelated to its enzyme activity[23] (see Section 5.7). Among ureolytic prokaryotic organisms, this chapter will focus on human pathogenic bacteria showing ureolytic activity. The most common sites for ureolytic bacteria infections in humans are urinary and gastrointestinal tracts.[24,25] The pathogenesis of the bacteria is strongly related to the ureolytic activity, in particular the pH increase and the toxicity of released ammonia. For example, *Proteus mirabilis* is a ureolytic bacterium that infects urinary and gastrointestinal tracts in human and animals,[26] causing kidney stone formation, catheter encrustation, pyelonephritis, ammonia encephalopathy, and hepatic coma.[2,27] Other bacteria involved in stone formation belong to *Pseudomonas, Klebsiella*, and

Staphylococcus spp.[24–28] *Helicobacter pylori* is probably the most important ureolytic human pathogen bacterium, acting as the major cause of several pathologies, such as gastritis, gastroduodenal ulcers, and cancer induced by chronic gastroduodenal infections.[29] It is a neutrophil bacterium able to survive in, and colonize, the acidic environment of the stomach by exploiting the pH increase caused by urease activity. Other lethal ureolytic microorganisms able to infect human body are *Mycobacterium tuberculosis*,[30] *Yersinia enterocolitica*,[31] and *Cryptococcus neoformans*.[32]

Urease activity is also widespread in soils, deriving primarily from extracellular immobilized soil urease, a residual of dead organisms.[33,34] The presence of urease in soils allows the worldwide use of urea as a nitrogen fertilizer,[35] because it accelerates the conversion of organic nitrogen into ammonium, a mineral nitrogen source usable by plants. However, the efficiency of soil nitrogen fertilization with urea is severely decreased by the activity of urease itself, when the latter is too fast. In fact, rapid urea breakdown may result in a decrease of soil fertility by unproductive loss of soil nitrogen as ammonia, which is released in the atmosphere, further contributing to environmental pollution. In addition, the release of ammonia can be toxic to plants, while the increase in soil alkalinity following urea hydrolysis causes damage to plantlets and seedlings, causing significant environmental and economic problems.[36–38]

The broad range of biological aspects in which urease is involved implies that a tight control of urease activity is required to counteract its negative effects, both in human health and in agriculture. For this purpose, several classes of compounds have been studied and tested as urease inhibitors in recent decades.[12,14,39] An integrated analysis of the main urease inhibitor classes will be discussed in Section 5.5.

5.3 Enzymology

The reaction of urea hydrolysis catalyzed by urease occurs in two steps: the first, strictly enzymatic, consists of the hydrolysis of urea to give ammonia and carbamate, followed by the spontaneous decomposition of carbamate to give a second molecule of ammonia and bicarbonate (Scheme 5.1).[3]

The reaction follows a typical Michaelis–Menten kinetic behaviour, with values of K_M in the 0.2–32 mM range and largely independent of pH.[12] By contrast, k_{cat} and consequently k_{cat}/K_M strongly depend on pH.[12] Urease is active in the 5–9 pH range. Several pH-dependent studies of urease

Scheme 5.1

activity have been carried out in order to elucidate the role of different conserved residues in the vicinity of the active site; they identified the presence of general acidic groups with pK_as of *ca.* 5 and *ca.* 6.9 and general basic groups with pK_as of *ca.* 6.3 and *ca.* 9, whose identities have not been fully elucidated yet.[40–42] Unlike the majority, a distinct subgroup of ureases possess an optimum pH in the range 2.0–4.5.[12] The overall amino acid sequence for these acidic enzymes does not differ from that of the neutral ureases as far as the active site residues are concerned, and the factors that determine this difference in the pH-dependent activity profiles are still unknown.

For ureases, weak substrate and product inhibition are observed. In particular, JBU shows uncompetitive inhibition by urea (K_i *ca.* 1–6 M) and noncompetitive inhibition by ammonium ions (K_i = 2–118 mM).[12]

Besides urea, urease catalyzes the hydrolysis of several urea analogues, such as formamide, acetamide, *N*-methylurea, *N*-hydroxyurea, *N,N′*-dihydroxyurea, semicarbazide, thiourea, as well as different kinds of phosphoric acid amides and esters.[43] Values of k_{cat} for these substrates are *ca.* 2–3 orders of magnitude lower than that observed for urea, and a kinetic characterization is complicated due to their simultaneous role as enzyme inhibitors (Section 5.5).[12]

5.4 Urease Structures

Initial structural information on the active site structure in urease was obtained by a long list of spectroscopic studies on JBU. The presence of six-coordinate octahedral Ni(II) in the JBU active site was first revealed by optical absorption spectroscopy,[7,44,45] and later confirmed by X-ray absorption spectroscopy, which further suggested the presence of Ni(II) ions coordinated to three histidine N atoms at 2.04 Å, two O atoms at 2.07 Å, and one O atom at 2.25 Å.[46–48] Magnetic susceptibility experiments on JBU indicated the presence of weakly anti-ferromagnetic coupling (J = −6.3 cm^{-1}) between two high spin (S = 1) and closely spaced octahedral Ni(II) ions.[49] This conclusion received further support from later EXAFS studies, which indicated the appearance of a new peak in the Fourier transform upon addition of β-mercaptoethanol (BME) to JBU that could be fitted using a model that involved the presence of two Ni(II) ions separated by 3.26 Å and bridged by the thiolate group of BME.[50] In the case of bacterial ureases, X-ray absorption spectroscopy studies carried out on urease from *Klebsiella aerogenes* (KAU)[50] and *Sporosarcina pasteurii* (formerly known as *Bacillus pasteurii*, SPU hereafter)[51] provided a picture of the active site essentially identical to that of JBU.

A more detailed characterization of the urease active site and its overall three-dimensional structure was made available with the first report of the crystal structure of KAU, in 1995.[52] Nowadays, the total number of full structures of ureases in the Protein Data Bank amounts to 49 (Table 5.1). Among these, 44 are ureases from bacterial sources, 4 are from jack bean, and one is from pigeon pea (*Cajanus cajan*).

Table 5.1　Selected structural parameters for urease structures currently available in the PDB. Adapted with permission from ref. 14. Copyright (2014) American Chemical Society.

PDB code	Enzyme source	Enzyme form	pH^a	T^b (K)	Metal, ligands^c	Flap status	Resolution (Å)	Reference
2KAU	*Klebsiella aerogenes*	Wild type	7.5	298	2Ni	Closed	2.20	52
1KRA	*K. aerogenes*	Wild type	7.5	298	—	Closed	2.30	59
1FWJ	*K. aerogenes*	Wild type	7.5	298	2Ni, water	Closed	2.20	65
1EJX	*K. aerogenes*	Wild type	7.5	100	2Ni, water	Disordered	1.60	Unpublished
1EJW	*K. aerogenes*	Wild type	7.5	298	2Ni, water	Closed	1.90	Unpublished
1A5K	*K. aerogenes*	αLys217Glu	7.5	298	—	Closed	2.20	66
1A5M	*K. aerogenes*	αLys217Ala	7.5	298	—	Closed	2.00	66
1A5N	*K. aerogenes*	αLys217Ala	7.2	298	2Ni, formate	Disordered	2.40	66
1A5L	*K. aerogenes*	αLys217Cys αCys219Ala	7.5	298	—	Disordered	2.20	66
1A5O	*K. aerogenes*	αLys217Cys αCys219Ala	7.2	298	2Ni, formate	Closed	2.50	66
1FWA	*K. aerogenes*	αCys319Ala	7.5	298	2Ni, water	Closed	2.00	65
1FWB	*K. aerogenes*	αCys319Ala	6.5	298	2Ni, water	Closed	2.00	65
1FWC	*K. aerogenes*	αCys319Ala	8.5	298	2Ni, water	Closed	2.00	65
1FWD	*K. aerogenes*	αCys319Ala	9.4	298	2Ni, water	Closed	2.00	65
1FWE	*K. aerogenes*	αCys319Ala	7.5	298	2Ni, AHA	Disordered	2.00	65
1FWF	*K. aerogenes*	αCys319Asp	7.5	298	2Ni, water	Disordered	2.00	65
1FWG	*K. aerogenes*	αCys319Ser	7.5	298	2Ni, water	Closed	2.00	65
1FWH	*K. aerogenes*	αCys319Tyr	7.5	298	2Ni, water	Open	2.00	65
1EJS	*K. aerogenes*	αHis219Asn	7.5	298	2Ni, water	Closed	2.00	124
1EJT	*K. aerogenes*	αHis219Gln	7.5	298	2Ni, water	Closed	2.00	124
1KRB	*K. aerogenes*	αHis219Ala	7.5	298	2Ni, water	Closed	2.50	59
1KRC	*K. aerogenes*	αHis320Ala	7.5	298	2Ni, water	Closed	2.50	59
1EJU	*K. aerogenes*	αHis320Asn	7.5	298	2Ni, water	Disordered	2.00	124
1EJV	*K. aerogenes*	αHis320Gln	7.5	298	2Ni, water	Disordered	2.40	124
1FWI	*K. aerogenes*	αHis134Ala	7.5	298	1Ni, water	Disordered	2.00	61
1EJR	*K. aerogenes*	αAsp221Ala	7.5	298	2Ni, water	Disordered	2.00	124
1EF2	*K. aerogenes*	Wild type	7.5	298	2Mn, water	Closed	2.50	129

			pH[a]	Temp[b]				Ref.
2UBP	Sporosarcina pasteurii	Wild type	6.8	100	2Ni, water	Open	2.00	53
1UBP	S. pasteurii	Wild type	6.8	100	2Ni, BME	Open	1.65	71
3UBP	S. pasteurii	Wild type	6.8	100	2Ni, DAP	Closed	2.00	53
4UBP	S. pasteurii	Wild type	6.8	100	2Ni, AHA	Open	1.55	77
1IE7	S. pasteurii	Wild type	6.8	100	2Ni, PHO	Open	1.85	90
1S3T	S. pasteurii	Wild type	6.8	100	2Ni, H_3BO_3	Open	2.10	91
4AC7	S. pasteurii	Wild type	6.8	100	2Ni, citrate	Open	1.50	94
4CEU	S. pasteurii	Wild type	6.5	100	2Ni, water	Open	1.58	97
4CEX	S. pasteurii	Wild type	6.5	100	2Ni, fluoride	Open	1.59	97
5A6T	S. pasteurii	Wild type	6.5	100	2Ni, sulfite	Open	1.65	74
5FSE	S. pasteurii	Wild type	6.5	100	2Ni, 1,4BQ	Open	2.07	111
5FSD	S. pasteurii	Wild type	6.5	100	2Ni, 2,5BQSO3	Open	1.75	111
5G4H	S. pasteurii	Wild type	6.5	100	2Ni, catechol	Open	1.50	119
1E9Z	Helicobacter pylori	Wild type	6.5	100	2Ni, water	Closed	3.00	54
1E9Y	H. pylori	Wild type	6.5	100	2Ni, AHA	Open	3.00	54
3QGA	H. mustelae	Wild type	7.4	100	2Fe	Disordered	3.00	131
3LA4	Canavalia ensiformis	Wild type	8.8	100	2Ni, PHO	Open	2.05	56
4H9M	C. ensiformis	Wild type	N.a.[d]	N.a.	2Ni, AHA	Open	1.52	Unpublished
4GY7	C. ensiformis	Wild type	N.a.	N.a.	2Ni, PHO	Open	1.49	Unpublished
4GOA	C. ensiformis	Wild type	N.a.	N.a.	2Ni, fluoride	Open	2.20	Unpublished
4G7E	C. cajan	Wild type	N.a.	N.a.	2Ni, water	Open	2.20	182
4Z42	Yersinia enterocolitica	Wild type	7.5	100	2Ni, water	Disordered	3.01	Unpublished

[a]pH of crystallization.
[b]Temperature for data collection.
[c]AHA: acetohydroxamic acid; BME: β-mercaptoethanol; DAP: diamidophosphate; PHO: phosphate.
[d]Not available.

Four structures of KAU refer to the native state of the recombinant wild-type enzyme, at different levels of refinement or data collection temperature (PDB codes 2KAU, 1FWJ, 1EJX, and 1EJW), while one structure involves the apo-form (PDB code 1KRA) and another one the manganese-substituted enzyme (PDB code 1EF2). In addition, 21 structures of KAU mutants are available. In the case of SPU, two structures involve the resting state of the enzyme (PDB codes 2UBP and 4CEU), while eleven additional structures report on complexes with inhibitors. Two structures of *H. pylori* urease (HPU, PDB codes 1E9Z and 1E9Y) are available, with significantly lower resolution than for the structures of KAU and SPU: one structure refers to the resting state (PDB code 1E9Z) and the other to the acetohydroxamic acid (AHA) complex (PDB code 1E9Y). The only available structure of the iron-containing alternative urease from *Helicobacter mustelae* (PDB code 3QGA) does not show any modeled electron density in the active site besides the metal ions and ligated residues. The most recent structure from a bacterial source is the native state urease from *Y. enterocolitica* (PDB code 4Z42). Five structures are available for plant urease, four of which are from JBU, complexed with AHA (PDB code 4H9M), phosphate (PHO) (PDB code 4GY7) or fluoride (PDB code 4GOA), and one from pigeon pea (PPU) in the native state (PDB code 4G7E). Figure 5.1 shows the protein architecture for KAU, SPU, HPU, and JBU.

KAU and SPU are representative of most bacterial ureases, with a quaternary structure composed of a trimer of trimers of the type $(\alpha\beta\gamma)_3$, with α, β, and γ being three different subunits.[52,53] The active site is found in the α subunit, giving rise to three active sites per biological unit. In other bacterial ureases, the quaternary structure is made of only two subunits to form $(\alpha\beta)_3$ trimers: the α subunit is highly homologous to that found in KAU and SPU, and the β subunit is a fusion peptide made of the β and γ subunits found in the KAU and SPU urease. In the case of HPU, four $(\alpha\beta)_3$ trimers form the spherically shaped tetramer of trimers $[(\alpha\beta)_3]_4$, containing twelve active sites[54] that correspond to the shape of the enzyme estimated using electron microscopy.[55] Finally, plant ureases are generally made up of a dimer of homotrimers $(\alpha_3)_2$, where the α subunit is derived from the fusion of the corresponding $\alpha\beta\gamma$ subunits found in bacteria.[56]

In all cases, the secondary and tertiary structures of ureases are very similar. The α subunits consist of a TIM barrel domain and a β-sheet domain, the β subunits are located on the external surface of the trimer and are mainly composed of β-sheets, and the γ subunits consist of domains containing both α-helices and β-sheets. An important structural feature of all ureases is the presence of a mobile helix-turn-helix motif in the α-subunit, flanking the active site cavity. The amino acid sequence of this motif is highly conserved among ureases from different sources. This flap has been found in an open or closed conformation, and is thought to be important in modulating the afflux of substrate and the efflux of products to and from the active site during catalysis, respectively, as well as moving the catalytically essential and conserved αCys322 and αHis323 residues (SPU numbering) by about 5 Å to and from the active site metal center (Figure 5.2).
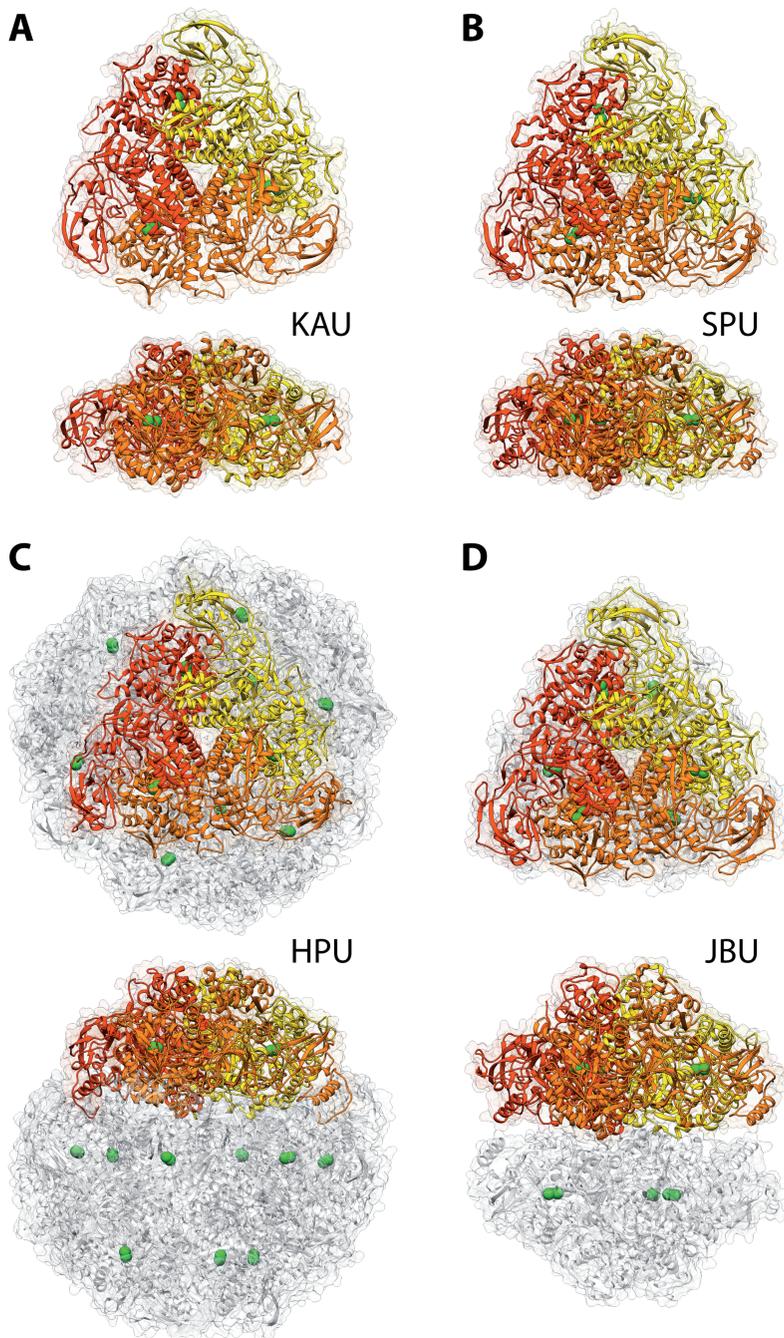
Figure 5.1 Ribbon diagram and solvent excluded surface of urease from (A) *K. aerogenes* (PDB code: 1EJZ), (B) *S. pasteurii* (PDB code: 4CEU), (C) *H. pylori* (PDB code: 1E9Z), and (D) jack bean (PDB code: 3LA4). Ribbon colors evidence the chains composing the trimer of oligomers (monomers in the case of jack bean) that constitute the minimal quaternary structure of urease. Ni(II) ions are shown as green spheres. The bottom panels are rotated by 90° around the horizontal axis *versus* the top panels. Adapted with permission from ref. 14. Copyright (2014) American Chemical Society.
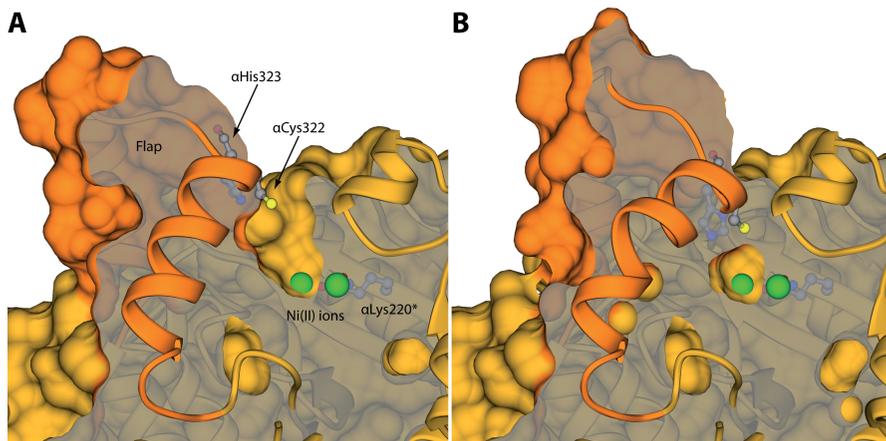
Figure 5.2 Ribbon diagram and longitudinal section of the open (A, PDB code 4CEU) and closed (B, PDB code 3UBP) conformations of the flexible flap in SPU highlighting the side chain of αLys220*, αCys322, and αHis323. The Ni(II) ions are shown as green spheres while other atoms are colored according to atom type.

A consensus has been reached on the structure of the coordination environment of the Ni(II) ions in the active site of urease. Indeed all structures so far determined indicate the presence of a dinuclear active site, with the two metal ions separated by 3.5–3.7 Å, bridged by the oxygen atoms of a carbamylated lysine residue, and bound to two histidines (Figure 5.3).

One Ni(II) ion is additionally bound to an aspartate carboxylate oxygen. The coordination geometry of each Ni(II) ion is completed by a terminally bound water molecule (W1 and W2) and by a nickel-bridging hydroxide ion (WB). The assignment of the protonation state for these solvent-derived nickel ligands was suggested[53] by the value of the first dissociation constant for $Ni(H_2O)_6^{2+}$ (pK_a = 10.6),[57] which supports the hypothesis that the terminally bound solvent molecules are neutral waters. Moreover, in water-bridged bimetallic complexes, the first pK_a for the bridging water decreases significantly to very acidic values, while the pK_a for hydroxide deprotonation is slightly lower than the pK_a of the first ionization of a water bound to a single Ni ion.[57] Therefore, the estimated pK_a for the deprotonation of the Ni bridging hydroxide (*ca.* 9–10) suggests that, at pH 8.0, the nickel-bridging solvent molecule is a hydroxide ion. Overall, in native urease two different types of ligands bridge the binuclear Ni cluster, the carboxylate group of the carbamylated lysine and the hydroxide ion, accounting for the observation of weak antiferromagnetic coupling.[49]

An extended network of hydrogen bonds stabilizes the nickel-bound solvent molecules in the active site: W1 is at 2.9 Å from αHis222 Nε (the following discussion follows the SPU residue numbering scheme), which is protonated and acts as a hydrogen-bonding donor, as deduced from the interaction of
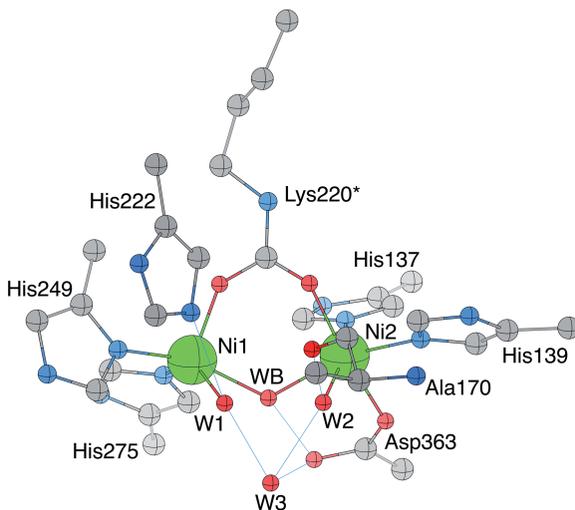
Figure 5.3 Crystallographic structural model of the *Sporosarcina pasteurii* urease active site in the resting state (PDB code 4CEU). The nickel ions are represented in green, while CPK coloring is used for all other atoms. Hydrogen bonds are shown as thin blue lines. The SPU residue-numbering scheme is used. Only selected residues are shown. Adapted with permission from ref. 14. Copyright (2014) American Chemical Society.

αHis222 Nδ with the peptide NH group of αAsp224 (at 2.9 Å). In contrast, W2 forms a hydrogen bond with αAla170 O (at 2.9 Å), which acts as hydrogen bond acceptor. This ligand arrangement yields one pentacoordinated Ni(II) ion (Ni1) with a distorted square-pyramidal geometry, and one Ni(II) ion (Ni2) hexacoordinated with a distorted octahedral geometry. An additional water molecule (W3) is part of a hydrogen-bonding network completing a tetrahedral cluster of four water/hydroxide molecules in close proximity to the Ni(II) ions, hinting at the existence of an active site cavity designed to stabilize a tetrahedral transition state and/or intermediate. In the structure of apo-KAU (PDB code 1KRA, Table 5.1), where the Ni(II) ions were chemically removed by lowering the pH in the presence of chelating agents, a treatment that decarbamylates αLys217, the positions of the remaining ligands are the same as in the holoenzyme, indicating a pre-organized metal binding site structure. The positions of conserved amino acid residues not involved in Ni binding but thought to be important in the catalytic mechanism (αAla170, αHis222, αGly280, αCys322, αHis323, αArg339, and αAla366) are largely invariant in SPU and KAU, except for αHis323, due to the different conformation of the active site flap. Some differences in the structure of active site residues are observed in the case of native HPU: in particular, the carbamylated lysine appears to bridge the two nickel ions using only one of the two terminal oxygen atoms, a consequence of an apparent rotation of the terminal $-NH-CO_2^-$ moiety along the N–C bond by about 90° with respect to the

position found in SPU and KAU. In addition, some distances between the nickel ions and the coordinating residues are very large (ranging between 2.6 and 2.7 Å) as compared to those found in KAU and SPU (2.0–2.1 Å), with an unusual orientation of the histidine imidazole rings around the metal ions, a very short (2.1 Å) Ni–Ni distance, and non-bridging hydroxide. These differences are most likely ascribed to the much lower resolution of this structure as compared to those of KAU and SPU (see Table 5.1).

All structural investigations on urease mutants have been carried out using KAU, and the residue-numbering scheme for this urease will be used to describe the attempts to relate their structures to the enzymatic mechanism. The structure of the αHis219Ala mutant (PDB code 1KRB), which features a much lower affinity for the substrate than the wild-type enzyme (K_M = 1100 *versus* 2.3 mM) and a *ca.* 30-fold decrease of k_{cat},[58] and thus indicates a role for αHis219 in substrate binding, shows structural identity with native KAU. The structure of the αHis320Ala mutant, (PDB code 1KRC)[59] shows an RMSD for Cα atoms of only 0.1 Å from wild-type KAU, and displays only a small change in K_M but a *ca.* 30 000-fold decrease in k_{cat} with respect to the wild-type enzyme,[60] while not showing the pK_a = 6.5 observed for wild-type urease.[58,60] The removal of one nickel histidine ligand in the catalytically inactive mutant αHis134Ala (ref. 58) resulted in an enzyme featuring only Ni1 in the active site (PDB code 1FWI),[61] with the position of all remaining active site residues, including the carbamylated lysine, invariant with respect to the wild-type enzyme structure. Chemical modification of αCys319, located on the flexible flap covering the active site of KAU, blocks enzyme activity,[62,63] while the αCys319Ala mutant is still *ca.* 50% as active as the wild type urease.[64] Structures of the αCys319Ala mutant were determined at pH 6.5 (PDB code 1FWB), pH 7.5 (PDB code 1FWA), pH 8.5 (PDB code 1FWC), and pH 9.4 (PDB code 1FWD),[65] with no significant structural differences observed between the structure of the wild-type and mutant enzymes except for a much reduced mobility of the flexible flap covering the active site in the mutant. The structures of αCys319Asp (PDB code 1FWF), αCys319Ser (PDB code 1FWG), and αCys319Tyr (PDB code 1FWH), which feature, respectively, 0.03%, 4.5%, and 0% of the activity observed for the wild type, also indicate the same active site environment but with a much higher mobility of the flap.[65] These evidences indicate that αCys319 (αCys322 in SPU), largely conserved in ureases, is somehow involved in catalysis, possibly having a role in positioning other key residues that are located on the mobile flap or in the active site, possibly the adjacent αHis320 (αHis323 in SPU), in a conformation best suited for catalysis. The structures of the αLys217Glu (PDB code 1A5K), αLys217Ala (PDB code 1A5M), and αLys217Cys/αCys319Ala (PDB code 1A5L) mutants of KAU, which were structurally characterized to establish the importance of the carbamylated lysine ligand for urease activity,[66] reveal the complete absence of bound Ni ions, indicating the need for the longer side chain in order to bind the nickel(II) ions in the active site defined by the four His ligands. The addition of formate to these mutants allowed for a chemical rescue of the enzyme, with the structures of the αLys217Ala–formate–Ni

complex (PDB code 1A5N) and αLys217Cys–formate–Ni complex (PDB code 1A5O) revealing the presence of a dinuclear Ni center bridged by formate instead of the carbamate group of αLys217 as in wild-type KAU.[66]

5.5 Urease Inhibitors

As briefly described in Section 5.2, several classes of compounds have been tested as urease inhibitors to counterbalance urease adverse effects both in medicine and the agro-environmental field.[2,12,14,27,39,67] In this section all urease inhibitor classes that have been characterized so far will be discussed.

5.5.1 Sulfur Compounds

Among sulfur compounds acting as urease inhibitors, thiols represent a class of historic importance, having been used for structural characterization of the active site of plant urease.[45,68] Thiols inhibit urease with a competitive mechanism in their thiolate anion form RS^-.[69] The X-ray structure of the SPU : BME complex (PDB code 1UBP, Figure 5.4A)[70,71] revealed the structural basis for the inhibition of urease by thiol compounds.
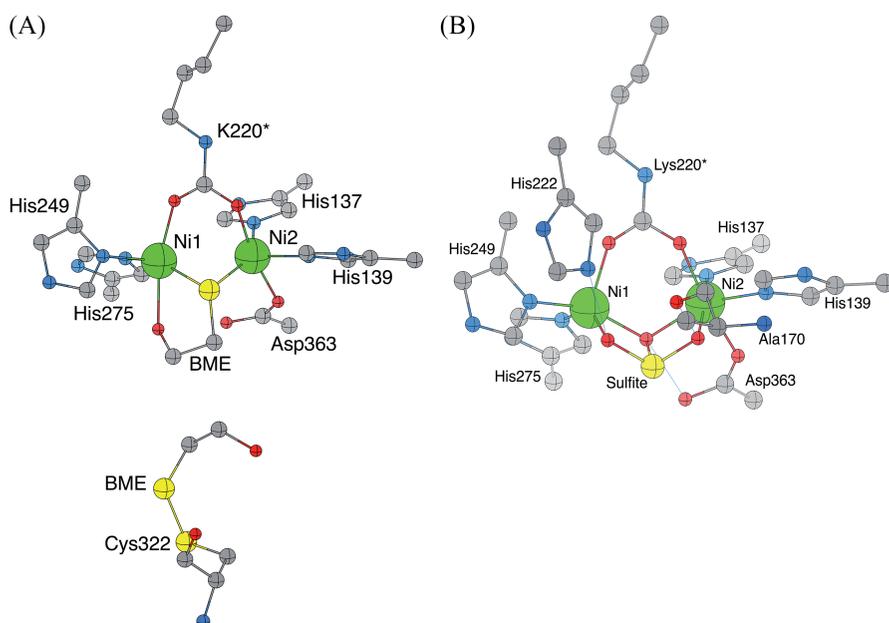
Figure 5.4 Crystallographic structural model of the *Sporosarcina pasteurii* urease active site in complex with (A) BME (PDB code 1UBP) and (B) sulfite (PDB code 5A6T). The nickel ions are represented in green, while CPK coloring is used for all other atoms. Hydrogen bonds are shown as thin blue lines. The SPU residue-numbering scheme is used. Only selected residues are shown. Part (A): adapted with permission from ref. 14. Copyright (2014) American Chemical Society.

The thiolate group of BME bridges the two Ni(II) ions and chelates the metal center with the alcoholic group interacting with Ni1, the interaction being further stabilized by a hydrogen bond with the carbonyl oxygen of the conserved αGly280. Additionally, a second molecule of BME forms a disulfide bond with the conserved αCys322 residue belonging to the mobile flap. The consequent formation of a hydrogen bond between the α-hydroxyl group belonging to BME and the carbonyl oxygen atom of αAla366 decreases the flexibility of the flap, plugging the entrance to the active site by steric hindrance.

Besides thiols, sulfite is a well-documented sulfur compound that acts as a competitive urease inhibitor[72,73] as well as a stabilizer of the Ni(II)–urease complex in certain cases.[43] Biochemical studies performed on SPU demonstrated that sulfite inhibition is pH-dependent in the 6.5–8.0 range of pH, decreasing with increasing pH and becoming negligible at pH 8.0.[74] The binding mode of sulfite on SPU was revealed by the X-ray structure of the SPU : sulfite complex (PDB code 5A6T, Figure 5.4B).[74] The inhibitor binds the two Ni(II) ions in the active site in a tridentate binding mode, using two oxygen atoms [Ni1–O(1) and Ni2–O(2), respectively], and placing its third oxygen atom, O(B), in the bridging position, thus replacing the terminal and the bridging water molecules that were bound to Ni1 and Ni2 in the native form of the enzyme. The SPU : sulfite interaction is also stabilized by a hydrogen bond network between ligand and residues that surround the active site cavity: sulfite oxygen O(1) forms a hydrogen bond with αHis222 Nε (at 2.65 Å), while two water molecules are H-bonded to O(1) at 3.19 Å and 3.21 Å. Sulfite oxygen O(2) is placed at 2.78 Å from the carbonyl backbone αAla170 O, suggesting a possible H-bond in which the sulfite O(2) could be protonated. Finally, the bridging sulfite O(B) is placed at 2.48 Å from αAsp363 Oδ2, again suggesting that these two atoms could share a hydrogen atom through an H-bond. The pH-dependent behavior of the inhibition, as well as the hydrogen bond network briefly described, exclude the inhibitor acting in the di-anionic SO_3^{2-} form, rather suggesting a bisulfite mono-protonated moiety (HSO_3^-) as the actual inhibitor, as previously reported.[73]

5.5.2 Hydroxamic Acids

Hydroxamic acids act as urease inhibitors[69,75,76] by exploiting their metal-binding properties.[39] Among them, AHA has been the most studied derivative and it acts as a slow-binding inhibitor for plant, bacterial, and fungal ureases.[12] X-Ray crystal structures of urease complexed with AHA are available for SPU (PDB code 4UBP),[77] HPU (PDB code 1E9Y),[54] and a mutant of KAU (PDB code 1FWE).[65] In all cases, AHA interacts with the two Ni(II) ions in the active site (Figure 5.5). The hydroxamate oxygen atom bridges the two Ni(II) ions, while the carbonyl oxygen of AHA chelates one nickel atom in a similar fashion as observed for BME (Figure 5.4A). Two H-bonds, between the carbonyl AHA O atom and αHis222 NεH, and between the AHA–NH group and the Oδ2 atom of αAsp363 residue, stabilize the SPU : AHA interaction.
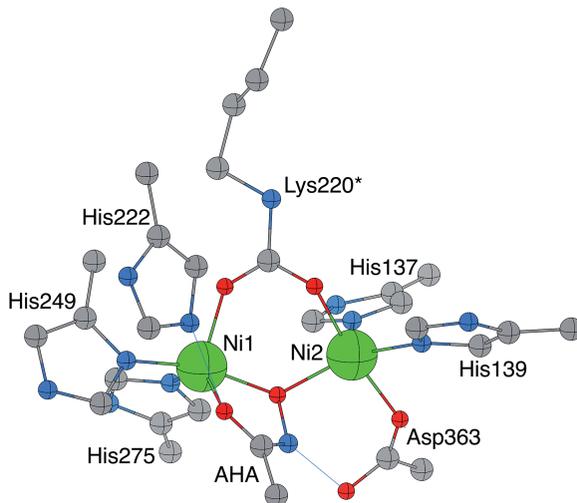
Figure 5.5 Crystallographic structural model of the *Sporosarcina pasteurii* urease active site in complex with AHA (PDB code 4UBP). The nickel ions are represented in green, while CPK coloring is used for all other atoms. Hydrogen bonds are shown as thin blue lines. The SPU residue-numbering scheme is used. Only selected residues are shown. Adapted with permission from ref. 14. Copyright (2014) American Chemical Society.

AHA has been commercialized for medical treatments of urinary tract infections caused by pathogenic bacteria.[24] However, although it is a urease inhibitor, AHA and other hydroxamic acid derivatives produce severe side effects.[39]

5.5.3 Phosphorus Compounds

Awareness of organic compounds containing phosphorus as inhibitory agents of urease dates back to the 1970s, when some organophosphorus insecticides were shown to inhibit soil urease.[78] In the same decade, Dixon and co-workers demonstrated the direct involvement of phosphoramidate in binding to the nickel ions in the JBU active site.[44] Nowadays, amide and ester derivatives of phosphoric and thiophosphoric acid are quite effective inhibitors of urease.[12] The latter are currently believed to be precursors that become inhibitors upon conversion into their oxygen analogues, even though direct crystallographic evidence for this hypothesis has not been obtained yet.[79,80] Owing to their inhibition properties, numerous derivatives have been intensively studied and patented for slowing down urea breakdown in soils and against infections by ureolytic bacteria.[39] Independent of the type of derivative, the initial enzymatic hydrolysis of the molecule generates diamidophosphate (DAP), which is proposed to act as the actual inhibitor.[81–83] Consistent with this hypothesis, the X-ray crystal structure of SPU crystallized in the presence of phenyl phosphorodiamidate (PPD) (PDB code 3UBP, Figure 5.6A)[53]

(A)   (B)
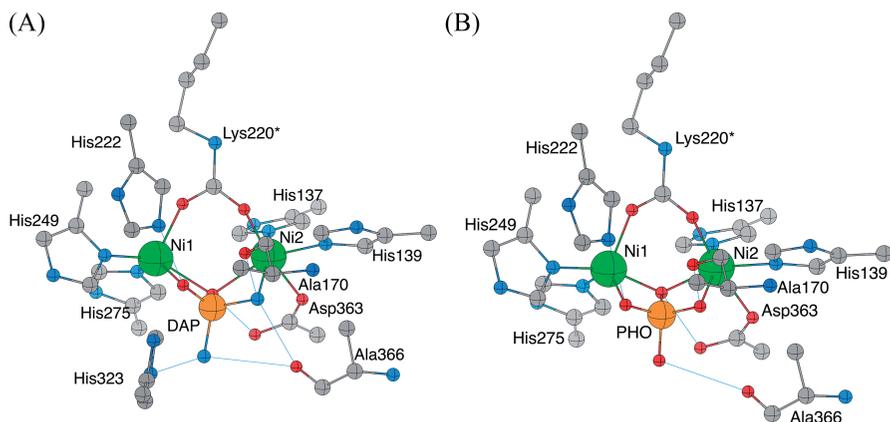
Figure 5.6 Crystallographic structural model of the *Sporosarcina pasteurii* urease active site in complex with (A) DAP (PDB code 3UBP) and (B) phosphate (PDB code 1IE7). The nickel ions are represented in green, while CPK coloring is used for all other atoms. Hydrogen bonds are shown as thin blue lines. The SPU residue-numbering scheme is used. Only selected residues are shown. Adapted with permission from ref. 14. Copyright (2014) American Chemical Society.

indeed revealed the presence of a molecule of DAP, produced *in situ* by the hydrolysis of PPD and the production of phenol.

DAP binds to Ni1 and to Ni2 through one oxygen and one nitrogen atom, respectively, where the Ni1-bound oxygen atom receives a hydrogen bond from αHis222 NεH, while the Ni2-bound DAP–NH$_2$ group forms two hydrogen bonds with the carbonyl oxygen atoms of αAla170 and αAla366. The second DAP oxygen bridges the two Ni(II) ions and is placed at hydrogen bonding distance from the αAsp363 Oδ2 atom. Finally, the second DAP nitrogen atom points away towards the cavity opening. The binding mode of the DAP molecule is considered to represent an analog of the transition state, replacing the tetrahedral cluster of four water/hydroxide molecules found in native SPU. This consideration is also strengthened by the closed conformation of the mobile flap that covers the active site cavity, which allows the formation of two additional hydrogen bonds between the distal DAP–NH$_2$ group and the αAla366 O and αHis323 Nε atoms.

Despite their strong inhibition properties, the main issue involving inhibitors based on the organophosphorus moiety is related to their low hydrolytic stability at acidic pH.[84,85] For this reason, some non-hydrolysable analogs of phosphoramidates, such as aminophosphinic acids, have been developed.[86,87]

Phosphate (PHO) has been found to act as a urease competitive inhibitor.[88] Its inhibition is pH-dependent in the pH range 5.0–8.0, becoming negligible at pH higher than 7.5–8.0.[69,89] A structural study of the SPU:PHO complex (PDB code 1IE7, Figure 5.6B)[90] shows a tetrahedral phosphate molecule binding to the binuclear Ni(II) active site through three atoms: a phosphate

oxygen bridges the two Ni(II) ions, and two oxygen atoms are bound terminally to each Ni center, with the fourth phosphate oxygen pointing away from the metal center. The pH-dependent behavior of urease inhibition by phosphate, which suggests interactions involving at least two protonation sites, with pK_as of *ca.* 7.2 and 6.5, as well as the H-bonding network established between the inhibitor and the active site residues, points out the actual protonation state of phosphate. The inhibitor is formally bound as the neutral phosphoric acid (H_3PO_4) in the pH range 5–6.5, with the pK_a observed at *ca.* 6.5 belonging to the production of the $H_2PO_4^-$ species in the pH range 6.5–7.2. Above this pH, loss of the second phosphate proton to form HPO_4^{2-} further destabilizes the enzyme–inhibitor interaction.[90]

5.5.4   Boric and Boronic Acids

Boric and boronic acids are competitive inhibitors of plant and bacterial ureases.[12] The crystal structure of the SPU : boric acid complex (PDB code 1S3T, Figure 5.7)[91] revealed that the inhibitor bridges the Ni(II) ions with two oxygen atoms, displacing three water molecules and leaving in place the bridging hydroxide.

The third inhibitor oxygen atom points away from the metal center. A detailed analysis of the H-bonding network surrounding the inhibitor established that the protonation state corresponds to the neutral $B(OH)_3$ molecule, confirming a previously formulated hypothesis.[69,92,93]
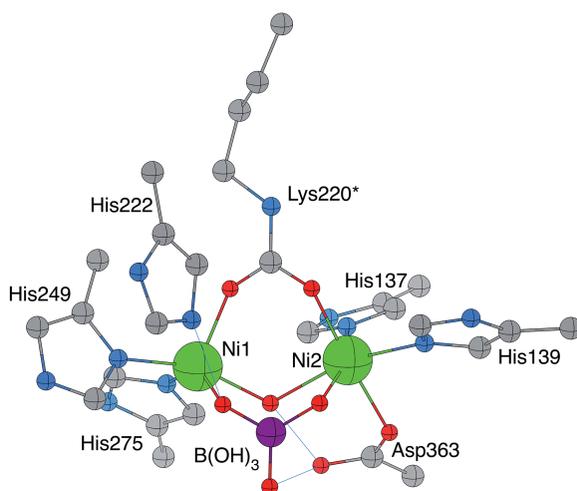
Figure 5.7   Crystallographic structural model of the *Sporosarcina pasteurii* urease active site in complex with boric acid (PDB code 1S3T). The nickel ions are represented in green, while CPK coloring is used for all other atoms. Hydrogen bonds are shown as thin blue lines. The SPU residue-numbering scheme is used. Only selected residues are shown. Adapted with permission from ref. 14. Copyright (2014) American Chemical Society.

5.5.5 Citrate

Citrate is a poor urease inhibitor at concentrations higher than 800 mM, while at lower concentrations (100–600 mM) the inhibition is masked by an activating effect.[94] The structure of the complex between citrate and SPU (Figure 5.8) shows the presence of a carboxylate group that binds the nickel ions in the active site, forming several hydrogen bonds involving the polar groups of the inhibitor and the surrounding residues, supposedly stabilizing this interaction that, considering the poor inhibition capability of citrate, is relatively weak.[94]

The rest of the active site residues, including the bridging solvent molecule WB, are structurally very similar to the native enzyme structure. The activating action of low concentrations of citrate and the inhibiting role at high concentrations suggested different binding sites of this species on the urease protein, with only the latter involving direct binding to the Ni(II) ions in the active site.[94]

5.5.6 Fluoride

Fluoride has been extensively studied as a urease inhibitor.[44,95,96] It carries out its inhibitory role through a complex mechanism, as revealed by biochemical studies performed on SPU.[97] In particular, fluoride exerts a mixed
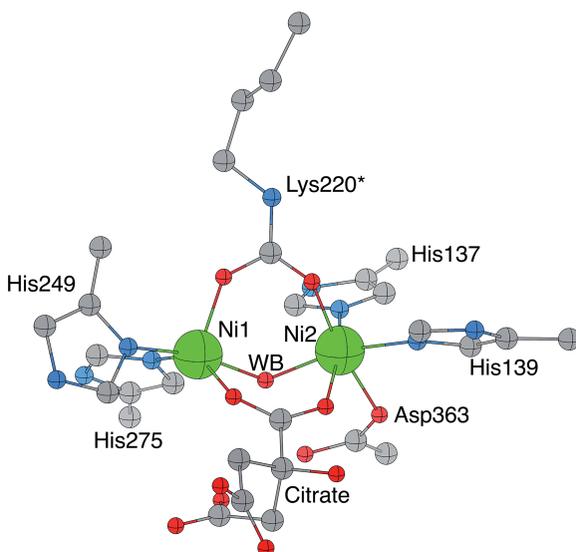
Figure 5.8 Crystallographic structural model of the *Sporosarcina pasteurii* urease active site in complex with citric acid (PDB code 4AC7). The nickel ions are represented in green, while CPK coloring is used for all other atoms. The SPU residue-numbering scheme is used. Only selected residues are shown. Adapted with permission from ref. 14. Copyright (2014) American Chemical Society.
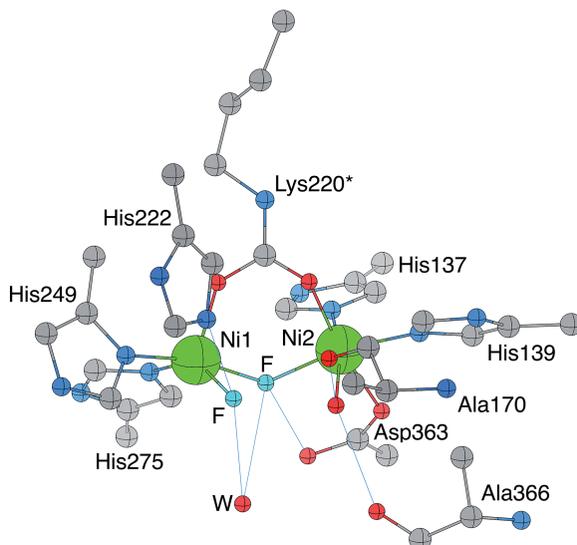
Figure 5.9 Crystallographic structural model of the *Sporosarcina pasteurii* urease active site in complex with fluoride (PDB code 4CEX). The nickel ions are represented in green, while CPK coloring is used for all other atoms. Hydrogen bonds are shown as thin blue lines. The SPU residue-numbering scheme is used. Only selected residues are shown.

competitive and uncompetitive inhibition. The pH dependence of the inhibition constants, investigated in the 6.5–8.0 range, reveals a predominant uncompetitive mechanism that is stronger at higher pH, and a lesser competitive inhibition that increases at lower pH. Five crystal structures of the enzyme in the native form and the same number of structures for the fluoride-inhibited urease were independently determined in order to distinguish fluoride from solvent molecules in the active site with solid statistic confidence.[97] The result of this detailed analysis indicates that one fluoride ion binds to Ni1 of the active site, while another fluoride ion replaces the nickel-bridging hydroxide (Figure 5.9).

In this way, the mixed inhibition mode of fluoride was interpreted as follows: the fluoride ion bound to Ni1 competes with the incoming substrate urea (more effective as the pH increases, due to the increase of the concentration of the fluoride anion), while the bridging fluoride ion replaces the bridging hydroxide (more effective as the pH decreases, which leads to protonation of the latter making it more prone to substitution), preventing it from carrying out the nucleophilic attack on the substrate urea.

5.5.7 Heavy Metals

Although structural data on urease–metal ion interactions are not available yet, biochemical information demonstrates that metal ions act as slow binding inhibitors on urease, with Hg(II), Ag(I), and Cu(II) ions being described

as the strongest.[12] Metal ions exert their inhibitory function in a concentration-dependent manner mainly by reacting with the sulfhydryl groups of cysteine residues, in particular with the conserved cysteine residue located on the mobile flap, whose modification results in the actual enzyme inhibition.[98] Additionally, Cu(II) and possibly Ag(I) can react with nitrogen- and oxygen-containing functional groups, contributing to an apparent inactivation process by protein aggregation.[98,99] In the case of Cu(II), generation of reactive oxygen species may determine the oxidation of further functional groups and subsequent enzymatic inactivation.[98]

Bismuth compounds represent a class of heavy metal containing molecules widely used as bactericidal agents in the treatment of peptic ulcers and *H. pylori* infections.[100–102] Although a detailed characterization of the bactericidal mode of action has not been elucidated, it has been demonstrated that Bi(III) ions inactivate urease by interacting with the thiol group of the mobile flap cysteine residue, at least contributing to bismuth inhibition of urease.[103]

5.5.8   Quinones

Quinones are a widely occurring class of molecules in nature, being involved in biological reduction–oxidation processes. They are also known for their bacteriostatic and fungicidal properties,[104,105] as well as their ability to inactivate proteins by either covalent modification of thiol groups through nucleophilic addition/substitution or quinone – inducing redox cycling, resulting in reactive oxygen species generation and indirect inactivation of protein functional groups.[106,107] The relevance of quinones as urease inhibitors was pointed out by Bremner and co-workers in the 1970s, demonstrating the inhibition properties of *p*-benzoquinones on soil urease and reporting 1,4-benzoquinone (*para*-benzoquinone, PBQ) as the most promising inhibitor.[108–110] In the recent past, Krajewska and co-workers carried out several kinetic studies on the inhibition of JBU by quinones, demonstrating a general common feature in the inactivation behavior towards urease.[107] These molecules showed a slow-binding concentration-dependent inhibition mechanism, consistent with the covalent modification of the thiol groups of urease, and in particular the conserved cysteine residue on the active site mobile flap. The strength of the inactivation was found to be correlated with the electrophilicity of the quinone ring, with quinones that possess electron-donating substituents being weaker urease inhibitors, due to a decreasing electrophilicity of the quinone ring.[107] Additionally, in the case of 1,4-naphthoquinone (1,4-NQ), a time- and concentration-dependent inactivation of urease has been shown, a mode of action that was ascribed to simultaneous covalent modification of the conserved cysteine on the active site flap and indirect thiol oxidation caused by 1,4-NQ-catalyzed redox cycling.[107] The molecular structural details of the inactivation of urease by quinones have been established by the X-ray crystal structures of SPU inactivated by PBQ in the presence and in the absence of sulfite ions (PDB code 5FSD and 5FSE, respectively, Figure 5.10A,B).[111]

PBQ covalently binds to the Sγ atom of the αCys322 residue located on the mobile flap through one carbon atom from the quinone ring. The PBQ
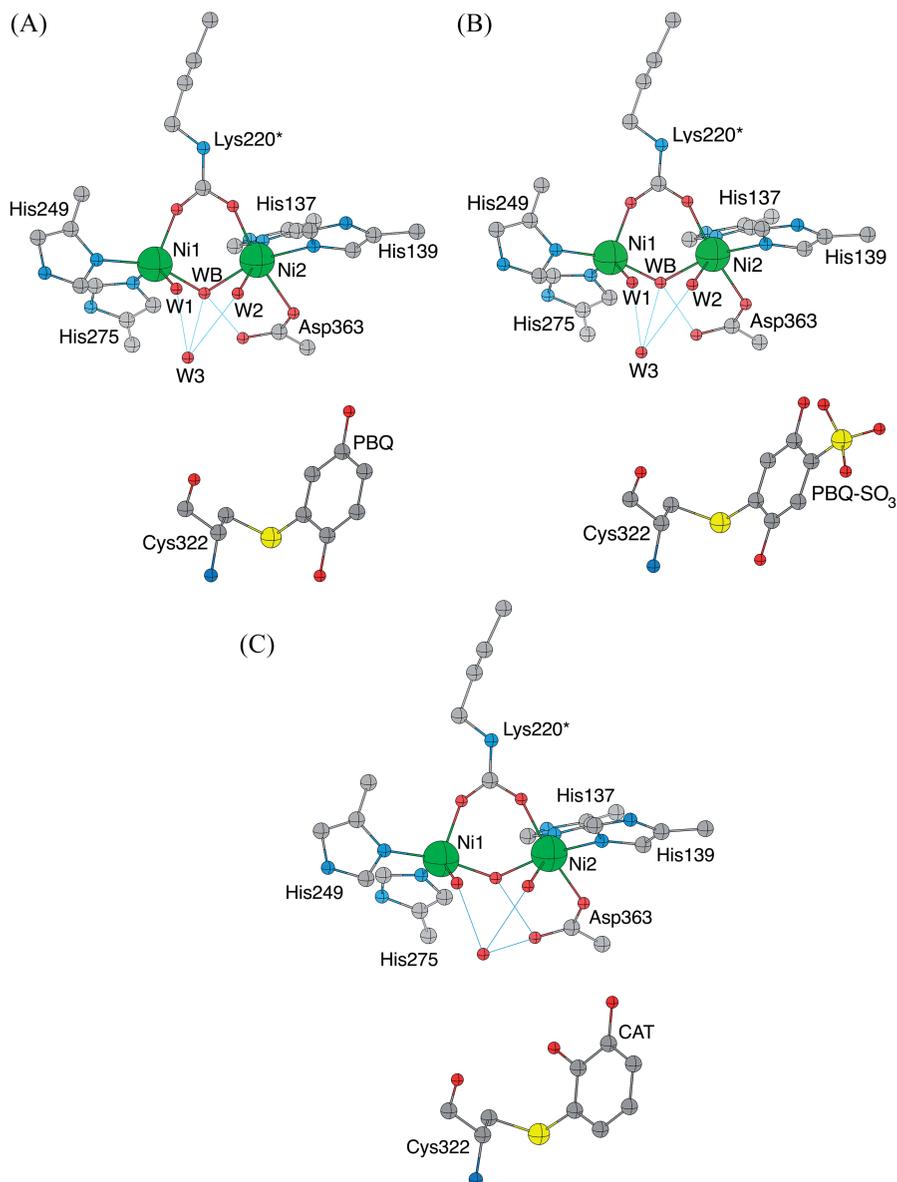
Figure 5.10 Crystallographic structural model of the *Sporosarcina pasteurii* urease active site in complex with (A) *p*-benzoquinone (PDB code 5FSE), (B) 2,5-dihydroxy-benzenesulfonate (PDB code 5FSD), and (C) catechol (PDB code 5G4H). The nickel ions are represented in green, while CPK coloring is used for all other atoms. Hydrogen bonds are shown as thin blue lines. The SPU residue-numbering scheme is used. Only selected residues are shown.

oxygen atom at the *meta* position with respect to the thiol group points towards the entrance of the active site channel, forming a hydrogen bond with a water molecule (at 2.85 Å) that is also stabilized by hydrogen bonds with the carbonyl carbon groups of αLeu365 (at 2.66 Å) and αLys169 (at 2.76 Å) residues (not shown in the figure). PBQ binding does not cause a closure of the channel within the mobile flap and the active site, suggesting that the inactivation of the enzyme by quinones is not caused by the blockage of substrate transport through the path leading into the active site, but it is rather due to the prevention of flap closure, a necessary event for the catalysis to occur. Interestingly, a site-directed mutagenesis experiment that replaced the same cysteine with tyrosine in KAU, supposedly having a similar effect on the active site entrance, also causes the abolition of catalytic activity.[64]

5.5.9   Polyphenols

Polyphenols are natural compounds generally recognized as beneficial to human health for their antioxidant properties. Several studies pointed out the efficacy of natural polyphenols on urease inhibition. Bremner and Douglas first demonstrated that catechol, the simplest molecule featuring a polyphenol scaffold, displayed inhibition properties towards soil urease.[108] In recent studies, the inhibition properties of several polyphenols and flavonoids against HPU have been examined, revealing that the two *ortho*-hydroxyl groups were essential for the inhibitory activity of polyphenol.[112–114] In addition, some polyphenols from green tea[115] have been recently shown to be HPU inhibitors. Kinetic studies on the inhibition of JBU by catechol showed an irreversible inhibition that takes place in a time- and concentration-dependent manner.[116] In that study, the authors claimed that inhibition could be due to the time-dependent oxidation of pyrocatechol to *ortho*-benzoquinone (OBQ), and increasing amounts could cause the increasing rate of urease inactivation.

The actual inhibition mechanism of catechol on urease is not clearly understood so far. The chemistry of polyphenols is complex.[117] They can be oxidized to form reactive *ortho*-semiquinone radicals and/or OBQs. The latter can undergo secondary reactions against protein functional groups, such as thiols and amines.[118] In addition, polyphenols can coordinate with transition metals to give catechol–metal complexes.[117] The recently reported crystal structure of SPU inhibited by catechol (PDB code 5G4H, Figure 5.10C)[119] reveals that the inhibitor covalently binds the Sγ atom of αCys322 residue, in a similar fashion as observed for PBQ.

5.6   Mechanism

In the absence of structural data about reaction intermediates that have short lifetimes, the available structures of urease in the resting hydrated state bound to boric acid (an analogue of the substrate) and with diamidophosphate (an analogue of the possible intermediate of the hydrolysis reaction) provided a basis for the proposal of a reaction mechanism for enzymatic urea hydrolysis (Figure 5.11).[120]
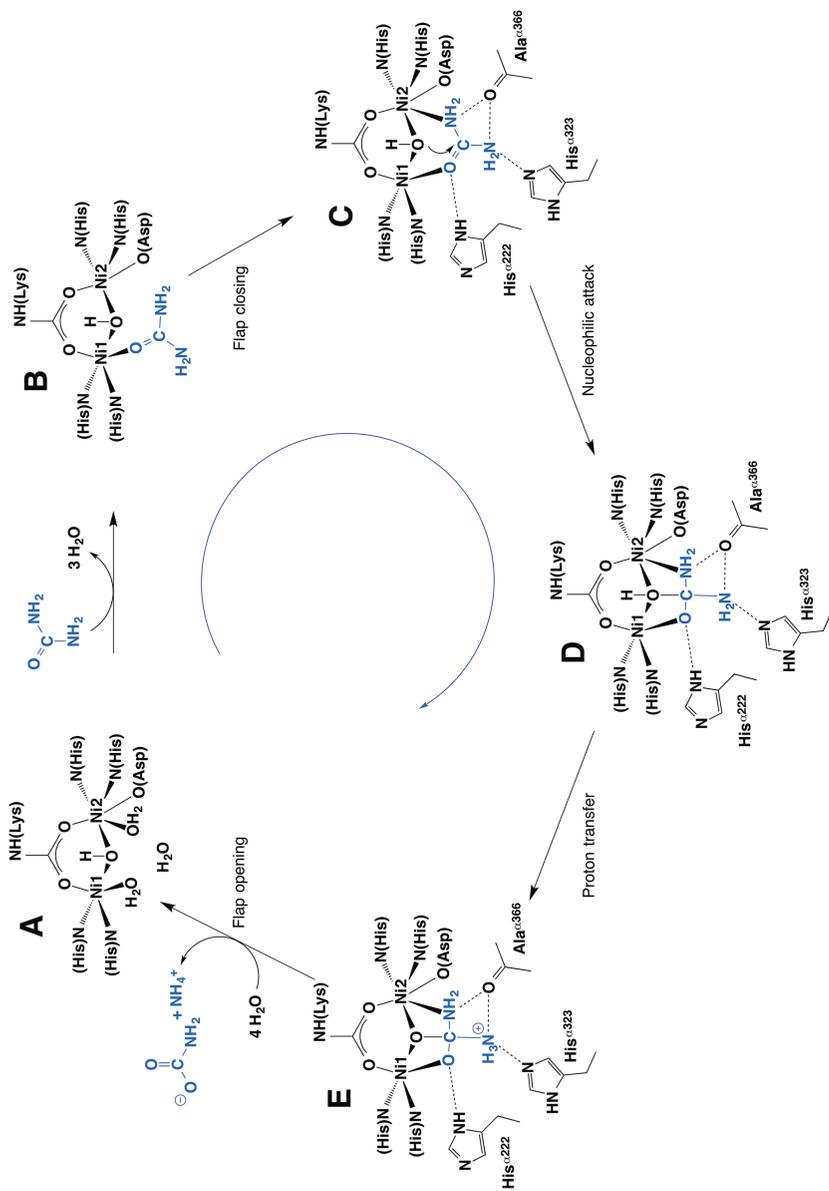
Figure 5.11 Structure-based urease catalytic mechanism of the enzymatic hydrolysis of urea. The SPU residue-numbering scheme is used. Adapted with permission from ref. 14. Copyright (2014) American Chemical Society.

In the following discussion, the SPU residue numbering scheme will be used. The presence of the tetrahedral water/hydroxide cluster in the proximity of the di-nickel center in the native enzyme (Figure 5.3), together with the binding mode of DAP (Figure 5.6A), suggests an orientation-specific mode of substrate binding in the enzyme active site that is designed to stabilize a tetrahedral transition state. The mechanism requires that upon urea entrance into the active site channel, with the flap in the open conformation, the structurally characterized hydrated active site of the resting enzyme evolves towards an initial substrate-bound intermediate where urea replaces the three water molecules bound to the Ni(II) ions. In particular, urea initially binds the more electrophilic and coordinatively unsaturated Ni1 using the urea carbonyl oxygen, an event that must be concomitant with the displacement of water molecules from the active site due to steric hindrance. The Ni1-bound urea O atom is stabilized in this position through the formation of an H-bond donated by the NH group of αHis222, analogous to what is observed in the structures of native SPU, as well as in SPU complexed with boric acid, diamidophosphate, phosphate, acetohydroxamic acid, and sulfite. This step is supported by docking and density-functional quantum chemistry calculations,[121] which also suggested that flap closure facilitates urea coordination to Ni2 *via* its –NH$_2$ group. This interaction is stabilized by a specific H-bonding network (H-bond donors on the urea carbonyl oxygen side and H-bond acceptors on the urea amide group side) that locates and steers the substrate to a precise orientation, thus allowing hydrolysis to occur. Urea is a poor chelating ligand because of the low Lewis base character of its –NH$_2$ groups. However, the formation of strong H-bonds with the nearby carbonyl oxygen atoms could enhance the basicity of the –NH$_2$ group and facilitate the interaction of the amide nitrogen with Ni2. The viability of a bidentate urea coordination mode is supported by the crystal structure of SPU in complex with B(OH)$_3$ (Figure 5.7). Boric acid has similar triangular shape and dimensions as urea, is isoelectronic with it, and has the same neutral charge, so it can be considered an inert structural analogue of the substrate. In the structure of urease complexed with B(OH)$_3$, two of the borate hydroxide moieties replace the water molecules terminally bound to Ni1 and Ni2, while the third borate hydroxide replaces the distal solvent molecule. This substrate-binding mode involves a direct role of both Ni(II) ions in binding and activating the substrate, therefore providing a rationale for the presence of a bimetallic active site in urease and explaining the lack of reactivity of urease containing a single nickel ion.[61] The presence of the Ni-bridging hydroxide in the complex of urease with B(OH)$_3$, placed at 2.1 Å from the B atom, in a direction almost perpendicular to the plane of the molecule, supports its role as the nucleophile attacking the carbon atom of urea to yield a tetrahedral transition state/intermediate. The bridging urea binding mode is the most efficient method to render the central carbon atom of urea electron poor and therefore prone to nucleophilic attack by the activated hydroxide. The kinetic inertia of a doubly coordinated nucleophile could be overcome by the weakening of the Ni–OH bonds upon substrate binding, as suggested by kinetic data on the inhibition

of KAU with fluoride,[95] by the structure and reactivity of inorganic models,[122] and by theoretical studies.[121] The high reactivity of the nickel-bridging hydroxide is also supported by the ability of the enzyme to hydrolyze PPD[81] and perhaps PHO, where the enzyme would simply induce an oxygen atom (or hydroxide) exchange on the PHO moiety. In this framework, only molecules able to react with the bridging hydroxide may bind the enzyme in a tridentate mode, as observed for DAP and PHO. This would explain why DAP strongly inhibits urease when it is formed by enzymatic hydrolysis of PPD, whereas it is a weak inhibitor if externally added to the native enzyme.[81] The nucleophilic attack by the Ni-bridging hydroxide onto the sp^2 carbon atom of urea yields a tetrahedral transition state containing a sp^3 carbon. The formation of a tetrahedral intermediate located between the two Ni(II) ions is supported by the structure of urease containing DAP in the active site (formed *in situ* from PPD, Figure 5.6A) that replaces the cluster of four water/hydroxide molecules.[53] The tetrahedral DAP bound to the di-nickel center represents an ideal transition state, or intermediate, analogue of the enzymatic reaction, while the structure of the PHO-inhibited urease (Figure 5.6B) supports the idea that the enzyme active site cavity is made to stabilize tetrahedral moieties. Closure of the flap would also be responsible for the stabilization of the catalytic transition state through the formation of multiple H-bonds with active site residues. The nucleophilic attack onto the Ni-bridging urea molecule profoundly modifies the electronic structure of the substrate and, in particular, increases the pK_a of the distal urea N atom not involved in Ni-binding, as supported by quantum mechanical calculations.[121] Furthermore, after the formation of the tetrahedral intermediate, the nickel-bridging –OH group, now part of a diamino(hydroxy)methanolate moiety, must now have a very low pK_a, and can therefore transfer the proton to the distal urea –NH$_2$ group, forming a C–NH$_3^+$ bond.[53] This proton transfer step could be mediated by αAsp363 Oδ2 through a dihedral rotation along the Cα–Cβ bond of the aspartate bound to Ni2, a movement observed to occur in the case of the SPU:AHA complex (Figure 5.5),[77] to bring this carboxylic oxygen atom, shown to be deprotonated,[123] close to the bridging hydroxide or, alternatively, to the distal –NH$_2$ urea group. The higher pK_a observed in the pH *versus* urease activity profile (*ca.* 9.5) corresponds to the estimated pK_a for the bridging hydroxide, suggesting that indeed the bridging hydroxide moiety must carry a proton for the enzyme to work properly.

By moving closer to the active site upon closure of the flap, the neutral imidazole side chain of the conserved αHis323 stabilizes the nascent C–NH$_3^+$ group.[53] Therefore, in this structure-based mechanism the bridging hydroxide acts as both the nucleophile and the general acid, while αHis323 stabilizes the positive charge that develops on the transition state acting as a H-bond acceptor. This scheme is supported by the structure of DAP-inhibited SPU (Figure 5.6A), in which the active site flap is in the closed conformation, allowing αHis323 to approach the Ni environment in the active site, and by the 10^3-fold reduction in k_{cat} upon mutation of αHi320 in KAU (corresponding to αHis323 in SPU).[124] The structure of the SPU:PHO (Figure 5.6B)

complex also reveals that another active site residue, αAla366, is important for the modulation of the protonation state of the distal urea –NH$_2$ group by adopting two different conformations, acting as a molecular switch able to provide stabilization for the protonation of the distal urea –NH$_2$ group by providing its backbone carbonyl group as a H-bond acceptor. The formation of the C–NH$_3^+$ moiety after proton transfer, and its stabilization by the catalytic histidine, causes the breakage of the distal C–N bond, with the subsequent release of ammonia.[121] The resulting carbamate spontaneously reacts with water, eventually forming another molecule of ammonia and hydrogen carbonate. The opening of the active site flexible flap facilitates the release of products, thus allowing bulk water to rehydrate the active site, regenerating the resting state of the enzyme. These steps could possibly occur in a concerted manner. This structure-based mechanism is in agreement with all available kinetic data, in particular with the pH-dependence of the enzyme activity and with the non-competitive inhibition by fluoride, which is suggested to replace the bridging hydroxide, and consistent with the hypothesis of the bridging hydroxide acting as nucleophile.

The presence of nickel as an enzyme metal cofactor in urease has intrigued the bioinorganic chemistry community since its discovery.[7] The structure-based mechanism discussed above could help to explain the requirement for Ni(II) ions instead of the more common and less toxic d^{10} closed-shell Zn(II) commonly observed in hydrolytic enzymes. The latter ion has a large positive charge density, which renders it able to act as a Lewis acid by polarizing substrates and preparing them for nucleophilic attack by hydroxide, and is resilient towards deleterious redox state changes. However, these properties are also applicable to Ni(II), which additionally features an open-shell d^8 electronic configuration that induces stereoelectronic restraints not available in the case of Zn(II). This property could be exploited by Ni(II) to drive the two substrates, urea and water, into the optimal spatial topology necessary for catalysis. Furthermore, Ni(II) has a higher affinity toward nitrogenous-based ligands than Zn(II),[125] thereby better stabilizing the binding of a urea –NH$_2$ group. Finally, Ni(II) ions possess multiple available binding sites due to their preference for an octahedral coordination sphere (as opposed to Zn(II), which mostly prefers tetrahedral[126]), thus facilitating both the bridging binding-mode of urea and stable metal ion binding to the protein through multiple sites for amino acids ligands.

Attempts to substitute the essential Ni(II) with other ions such as Zn(II), Co(II), and Mn(II) have been carried out to check the viability of a urease-based catalysis with non-native metals. Removal of both Ni(II) ions by treating JBU with EDTA at low pH causes irreversible inactivation of the protein.[127] Removal of a single Ni(II) ion from JBU was obtained by dialysis in the presence of citrate, and substitution of the more labile metal ion with Zn(II) or Co(II) produced catalytically inert urease.[128] This inactivation was also observed for KAU reconstituted with Zn(II), Co(II), and Cu(II), while in the case of the Mn-derivative *ca.* 0.3% of the activity was retained.[61] The crystal structure of the Mn-derivative of KAU was determined (PDB code 1EF2),

with active site features essentially identical to the native Ni-bound form.[129] Quite recently, an alternative urease in *Helicobacter mustelae* (HMU) has been reported, which is characterized by the absence of Ni(II), by inactivation in the presence of oxygen, and by Fe(II)-induced expression, observations that suggested the presence of Fe(II) ions in its active site.[130] This hypothesis was later confirmed, and the reported crystal structure showed little or no differences as compared to the Ni(II)-enzyme, while the catalytic efficiency appeared to be much smaller.[131] All together, these observations suggest that, in addition to the identity of the metal ions, the precise position and mobility of the metal ligands, as well as those of protein residues in the active site not involved in metal binding, are important in achieving optimal urease activity.

5.7 Non-Enzymatic Properties of Urease

Besides the well-established enzymatic aspects, non-enzymatic biological properties of ureases have also been established.[23] In general, plant and microbial ureases have been demonstrated to be toxic against filamentous fungi and yeasts by a mechanism involving cell wall and/or membrane permeabilization and consequent plasmolysis.[132,133] Furthermore, plant and some bacterial ureases have potent insecticidal effects, exhibiting neurotoxicity and affecting other physiological functions, such as diuresis, muscle contraction, and immunity.[134–137] Toxic effects of ureases have been characterized through *in vivo* and *in vitro* studies demonstrating them to be unrelated to enzymatic activity.[138–141] Although ureases are entomotoxic *per se*, studies conducted on canatoxin (CNTX), an isoform of JBU isolated from jack bean seeds[142,143] and soybean (*Glycine max*) embryo-specific urease (SBU), show that toxicity against insects derives, at least partly, from a small peptide (pepcanatox following canatoxin nomenclature) released from the ingested urease after its hydrolysis by insect cathepsin-like digestive enzymes. Recombinant analogues of these peptides, named Jaburetox and Soyuretox in the case of CNTX and SBU, respectively, also feature insecticidal and antifungal activity.[23,144] Jaburetox has been recently shown to be an intrinsically disordered polypeptide, possibly increasing its secondary structure content upon interaction with cell membranes.[145]

5.8 Microbial Induced Calcite Preparation by Ureolytic Bacteria

Calcium carbonate ($CaCO_3$) precipitation is a common (bio)geochemical process found in environments such as marine waters, freshwaters, and soils.[146,147] Different microorganisms are able to induce calcium carbonate precipitation (microbial induced carbonate precipitation, MICP) by creating an alkaline environment.[147,148] Among them, ureolytic microorganisms (*i.e. S. pasteurii*) are the most energy efficient.[149] The urease activity (production

of bicarbonate (HCO_3^-) and an overall increase of pH) in turn causes a shift of the bicarbonate equilibrium ($pK_{a(HCO_3^-)} = 10.33$ at 25 °C) towards an increase of the carbonate moiety (CO_3^{2-}) that precipitates in the presence of calcium ions as calcium carbonate crystals. MICP has been studied for several multidisciplinary applications, such as cementation of sands to enhance bearing capacity and liquefaction resistance, soil erosion control, remediation of soil and groundwater impacted by metals and radioactive elements, and many others.[150]

5.9   Urease Maturation Process: The Role of Accessory Proteins

Urease is initially expressed as an inactive apo-protein that undergoes a post-translational maturation process yielding a fully active enzyme (Figure 5.12).

This activation path involves two peculiar steps: (i) lysine carbamylation coupled with GTP hydrolysis and (ii) Ni(II) ion delivery into the urease active site (Chapter 14). The urease maturation process is usually carried out by four accessory proteins: UreD (called UreH in *H. pylori*), UreF, UreG, and UreE.[13,151] Genomes of most ureolytic organisms show structural and accessory proteins being encoded by genes clustered in one or more operons, with variable gene order depending on the source.[13] Overall, this organization reflects the need for both structural and accessory functions to achieve a fully active enzyme and suggests that their expression mechanism is concomitantly regulated.[13,151] The "classical" model for urease maturation involves UreD, UreF, and UreG acting together as a molecular chaperone in a UreD:UreF:UreG supercomplex (UreDFG) that drives GTP hydrolysis and lysine carbamylation, while UreE acts as the metallochaperone that delivers Ni(II) ions to the complex formed by apo-urease and UreDFG.[13,151] Even though the functional role of the single accessory proteins is not fully elucidated and no consensus has been reached on the sequence of events during the formation of the urease:UreDFG complex, the following model has been proposed (Figure 5.12A). UreD is the first protein to interact with the apoenzyme, being able to directly bind to urease.[152,153] It is considered to act as a protein scaffold that allows the interaction of the other chaperones with the enzyme,[154,155] although some studies suggest that it also likely facilitates nickel insertion into the active site.[61] Once the urease–UreD complex is formed, UreF binds it by interacting with UreD, inducing a conformational rearrangement of the enzyme.[152] UreF acts as the binding site for UreG and it has been demonstrated to have a role in enhancing the urease activation process, possibly by modulating the GTPase activity of UreG through a direct protein–protein interaction[156] and promoting access to the active site of Ni(II) ions and the CO_2 needed for the lysine carbamylation event.[157,158] The carbamylation process is carried out upon GTP hydrolysis by UreG, possibly through the formation of the carbamylating agent carboxyphosphate in the presence of CO_2.[159] UreG is the first
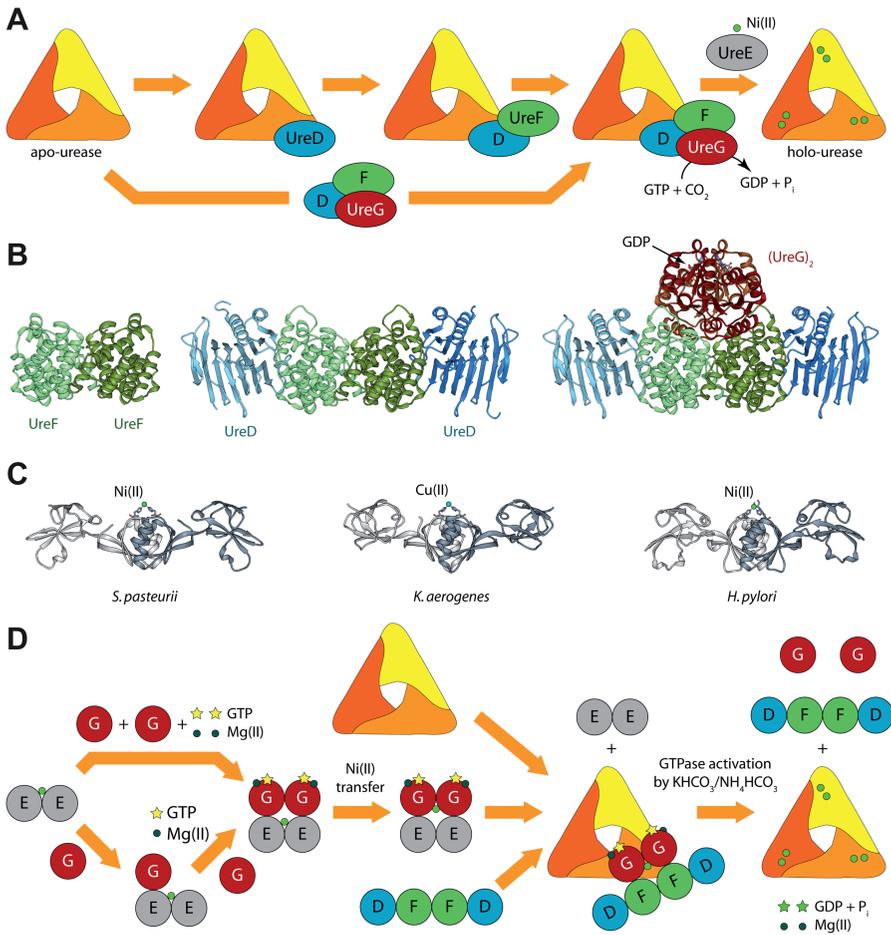
Figure 5.12 (A) Schematic representation of the "classic" proposed mechanisms for urease activation. (B) Ribbon diagram of the UreF (left-hand side) UreDF (center), and UreDFG (right-hand side) crystal structures (PDB codes 3CXN, 3SF5, and 4HI0, respectively). UreD, UreF, and UreG chains are colored in light blue, light green–yellow, and dark red, respectively, with the darker and lighter portions highlighting the different monomers. (C) Ribbon diagram of the UreE crystal structures from *Sporosarcina pasteurii* (left-hand side), *Klebsiella aerogenes* (center) and *Helicobacter pylori* (right-hand side). Chains are colored grey with the darker and lighter portions highlighting the different monomers. (D) Schematic representation of the new proposal for urease activation.[178]

case of an intrinsically disordered enzyme,[160] which can retain enzymatic activity owing to the rigidity of the active-site environment,[161] and is able to dimerize upon Zn(II) binding.[162,163] The macromolecular urease : UreDFG complex interacts with UreE, which acts as the metallochaperone that delivers Ni(II) to urease in order to complete the enzyme maturation (see Chapter 14 for a detailed discussion of nickel metallochaperones).[13,151] *In vivo* studies using yeast two-hybrid analysis[164,165] and co-immunoprecipitation assays,[165] as well as *in vitro* calorimetry and NMR spectroscopy,[166] indicated a direct interaction between UreE and UreG from *H. pylori*. The *in vitro* interaction between UreE and UreG has been observed also in *S. pasteurii*.[167]

Even though the structure of urease bound to any of the accessory proteins is not yet available, the crystal structure of the $(UreF)_2$ homodimer (PDB code 3CXN)[168] and the structure of the $(UreDF)_2$ (PDB code 3SF5)[169] and $(UreDFG)_2$ (PDB code 4HI0)[170] complexes from *H. pylori* have been recently reported (Figure 5.12B). The crystal structure of the $(UreDFG)_2$ complex contains two copies of each of UreF, UreD, and UreG, related by two-fold symmetry, forming a dimer of heterotrimers. Structural information on UreE proteins from various bacteria has been derived from numerous crystallographic studies: UreE from *S. pasteurii* (PDB codes 1EAR, 1EB0, and 4L3K),[171,172] *K. aerogenes* (PDB codes 1GMU, 1GMV, and 1GMW),[173] and *H. pylori* (PDB codes 3L9Z, 3LA0, 3NXZ, 3NY0, 3TJ8, 3TJ9, and 3TJA)[174,175] display a similar fold made by a symmetric homodimer (Figure 5.12C), with each monomer composed of two domains connected by flexible linkers.[176] UreE binds Ni(II) ions at the dimer interface by using one conserved histidine residue from each monomer and completing the coordination sphere by using histidine residues found in the C-terminal part of the chain.[166,172,175–177]

The entirety of the structural information from crystallography, together with UV-VIS spectroscopy, light scattering experiments, and GTPase activity assays performed on *H. pylori* UreG, recently suggested a new mechanism for the biosynthesis of the urease active site (Figure 5.12D).[178] In this new proposal, the Ni(II)-bound UreE dimer binds two apo-UreG monomers, facilitating GTP uptake by UreG in the presence of Mg(II) ions. The UreG binding to UreE can, in principle, occur either in a single or in a multistep process. In the $(UreEG)_2$ complex, the Ni(II) ion is then translocated from $(UreE)_2$ to $(UreG)_2$. Subsequently, the pre-formed $(UreDF)_2$ complex competes with $(UreE)_2$ for the $(UreG)_2$: Ni(II) complex to form the supercomplex apo-urease : $(UreDFG)_2$: Ni(II). Finally, the GTP hydrolysis performed by UreG is catalyzed by $KHCO_3/NH_4HCO_3$ to complete the nickel insertion into the apo-urease. Interestingly, the analysis of the *H. pylori* UreDFG structure highlighted the presence of a large cavity at the interface between UreF and UreG. The internal cavity contains several inner water molecules interconnected through a network aligned along the horizontal axis of the $UreD–UreF_2–UreD$ portion of the structure.[158] A deeper investigation of the complex structure revealed the presence of two nearly identical and symmetric tunnels going from the central cavity in the complex and exiting near the UreD C-terminus, passing through UreF and UreD. It has been hypothesized that Ni(II)

ions can proceed through these tunnels to reach the apo-urease active site. The importance of the tunnel passing through UreD has been recently highlighted in a work conducted on UreD from *K. aerogenes*.[179]

Some exceptions in urease maturation systems have been pointed out. Plants appear to lack homologs to UreE.[19] On the other hand, plant UreGs possess an extended nickel-binding nitrogen terminus rich in histidine and aspartate residues and it has been suggested to exploit this feature to replace UreE during Ni(II) delivery.[19] The genome of *Helicobacter mustelae*, a gastric pathogen of ferrets that additionally produces a Fe(II)-urease, contains two urease clusters.[180] The genome of *Bacillus subtilis* contains only the structural urease genes; nevertheless, this bacterium can synthesize an active nickel-containing urease, although with poor efficiency.[181]

5.10   Conclusions

The chemistry of the essential Ni(II) ions in the active site of urease – investigated by several research groups using different techniques, mainly X-ray crystallography, X-ray absorption spectroscopy, and enzymatic assays based on optical spectroscopy or calorimetry – has been revealed, and the peculiarities of this metal ion as compared to possible alternatives have been elucidated. This chapter thus represents the starting point from which the design and development of new and more efficient urease inhibitors for the control of urease activity in medical and agricultural applications will hopefully be achieved in the near future.

Acknowledgements

References

1. R. P. Hausinger, *Microbiol. Rev.*, 1987, **51**, 22–42.
2. H. L. T. Mobley and R. P. Hausinger, *Microbiol. Rev.*, 1989, **53**, 85–108.
3. R. L. Blakeley, J. A. Hinds, H. E. Kunze, E. C. Webb and B. Zerner, *Biochemistry*, 1969, **8**, 1991–2000.
4. N. E. Dixon, P. W. Riddles, C. Gazzola, R. L. Blakeley and B. Zerner, *Can. J. Biochem.*, 1980, **58**, 1335–1345.

5. C. Follmer, *Phytochemistry*, 2008, **69**, 18–28.

6. J. B. Sumner, *J. Biol. Chem.*, 1926, **69**, 435–441.

7. N. E. Dixon, C. Gazzola, R. Blakeley and B. Zerner, *J. Am. Chem. Soc.*, 1975, **97**, 4131–4132.

8. P. A. Karplus, M. A. Pearson and R. P. Hausinger, *Acc. Chem. Res.*, 1997, **30**, 330–337.

9. S. Ciurli, S. Benini, W. R. Rypniewski, K. S. Wilson, S. Miletti and S. Mangani, *Coord. Chem. Rev.*, 1999, **190–192**, 331–355.

10. R. P. Hausinger and P. A. Karplus, in *Handbook of Metalloproteins*, ed. A. Messerschmidt, R. Huber, T. Poulos and K. Wieghardt, John Wiley & Sons Ltd, Chichester, 2001, pp. 867–879.

11. S. Ciurli and S. Mangani, in *Handbook on Metalloproteins*, ed. I. Bertini, A. Sigel and H. Sigel, Marcel Dekker, New York, 2001, pp. 669–708.

12. B. Krajewska, *J. Mol. Catal. B: Enzym.*, 2009, **59**, 9–21.

13. B. Zambelli, F. Musiani, S. Benini and S. Ciurli, *Acc. Chem. Res.*, 2011, **44**, 520–530.

14. M. J. Maroney and S. Ciurli, *Chem. Rev.*, 2014, **114**, 4206–4228.

15. H. A. Krebs and K. Henseleit, *Hoppe-Seyler's Z. Physiol. Chem.*, 1932, **210**, 33–66.

16. W. H. R. Shaw and J. J. Bordeaux, *J. Am. Chem. Soc.*, 1955, **77**, 4729–4733.

17. R. L. Blakeley, A. Treston, R. K. Andrews and B. Zerner, *J. Am. Chem. Soc.*, 1982, **104**, 612–614.

18. B. P. Callahan, Y. Yuan and R. Wolfenden, *J. Am. Chem. Soc.*, 2005, **127**, 10828–10829.

19. C. P. Witte, *Plant Sci.*, 2011, **180**, 431–438.

20. L. E. Zonia, N. E. Stebbins and J. C. Polacco, *Plant Physiol.*, 1995, **107**, 1097–1103.

21. R. G. Winkler, D. G. Blevins, J. C. Polacco and D. D. Randall, *Trends Biochem. Sci.*, 1988, **13**, 97–100.

22. C. D. Todd, P. A. Tipton, D. G. Blevins, P. Piedras, M. Pineda and J. C. Polacco, *J. Exp. Bot.*, 2006, **57**, 5–12.

23. C. R. Carlini and R. Ligabue-Braun, *Toxicon*, 2016, **110**, 90–109.

24. R. A. Burne and Y.-Y. M. Chen, *Microbes Infect.*, 2000, **2**, 533–542.

25. C. M. Collins and S. E. D'Orazio, *Mol. Microbiol.*, 1993, **9**, 907–913.

26. B. D. Jones, C. V. Lockatell, D. E. Johnson, J. W. Warren and H. L. T. Mobley, *Infect. Immun.*, 1990, **58**, 1120–1123.

27. H. L. T. Mobley, M. D. Island and R. P. Hausinger, *Microbiol. Rev.*, 1995, **59**, 451–480.

28. I. J. Rosenstein and J. M. Hamilton-Miller, *Crit. Rev. Microbiol.*, 1984, **11**, 1–12.

29. C. Montecucco and R. Rappuoli, *Nat. Rev. Mol. Cell Biol.*, 2001, **2**, 457–466.

30. W. Lin, V. Mathys, E. L. Y. Ang, V. H. Q. Koh, J. M. Martínez Gómez, M. L. T. Ang, S. Z. Zainul Rahim, M. P. Tan, K. Pethe and S. Alonso, *Infect. Immun.*, 2012, **80**, 2771–2779.

31. G. M. Young, D. Amid and V. L. Miller, *J. Bacteriol.*, 1996, **178**, 6487–6495.

32. G. M. Cox, J. Mukherjee, G. T. Cole, A. Casadevall and J. R. Perfect, *Infect. Immun.*, 2000, **68**, 443–448.

33. N. M. Pettit, A. R. J. Smith, R. B. Freedman and R. G. Burns, *Soil Biol. Biochem.*, 1976, **8**, 479–484.

34. L. Gianfreda, A. De Cristofaro, M. A. Rao and A. Violante, *Soil Sci. Soc. Am. J.*, 1995, **59**, 811–815.

35. P. J. Stangel, in *Nitrogen in Crop Production*, ed. R. D. Hauck, American Society of Agronomy, Madison, WI, 1984, pp. 23–54.

36. J. M. Bremner and R. L. Mulvaney, in *Soil Enzymes*, ed. R. G. Burns, Academic Press, New York, 1978, pp. 149–196.

37. J. M. Bremner and M. J. Krogmeier, *Proc. Natl. Acad. Sci. U. S. A.*, 1989, **86**, 8185–8188.

38. J. M. Bremner, *Fert. Res.*, 1995, **42**, 321–329.

39. P. Kosikowska and Ł. Berlicki, *Expert Opin. Ther. Pat.*, 2011, **21**, 945–957.

40. S. B. Mulrooney, T. Zakharian, R. A. Schaller and R. P. Hausinger, *Arch. Biochem. Biophys.*, 2001, **394**, 280–282.

41. B. Krajewska and S. Ciurli, *Plant Physiol. Biochem.*, 2005, **43**, 651–658.

42. B. Krajewska, *J. Mol. Catal. B: Enzym.*, 2016, **124**, 70–76.

43. B. Zerner, *Bioorg. Chem.*, 1991, **19**, 116–131.

44. N. E. Dixon, R. L. Blakeley and B. Zerner, *Can. J. Biochem.*, 1980, **58**, 481–488.

45. R. L. Blakeley, N. E. Dixon and B. Zerner, *Biochim. Biophys. Acta*, 1983, **744**, 219–229.

46. S. S. Hasnain and B. Piggot, *Biochem. Biophys. Res. Commun.*, 1983, **112**, 279–283.

47. L. Alagna, S. S. Hasnain, B. Piggot and D. J. Williams, *Biochem. J.*, 1984, **220**, 591–595.

48. P. A. Clark, D. E. Wilcox and R. A. Scott, *Inorg. Chem.*, 1990, **29**, 579–581.

49. P. A. Clark and D. E. Wilcox, *Inorg. Chem.*, 1989, **28**, 1326–1333.

50. S. Wang, M. H. Lee, R. P. Hausinger, P. A. Clark, D. E. Wilcox and R. A. Scott, *Inorg. Chem.*, 1994, **33**, 1589–1593.

51. S. Benini, S. Ciurli, H. F. Nolting and S. Mangani, *Eur. J. Biochem.*, 1996, **239**, 61–66.

52. E. Jabri, M. B. Carr, R. P. Hausinger and P. A. Karplus, *Science*, 1995, **268**, 998–1004.

53. S. Benini, W. R. Rypniewski, K. S. Wilson, S. Miletti, S. Ciurli and S. Mangani, *Structure*, 1999, **7**, 205–216.

54. N.-C. Ha, S.-T. Oh, J. Y. Sung, K. A. Cha, M. H. Lee and B.-H. Oh, *Nat. Struct. Biol.*, 2001, **8**, 505–509.

55. J. W. Austin, P. Doig, M. Stewart and T. J. Trust, *J. Bacteriol.*, 1992, **174**, 7470–7473.

56. A. Balasubramanian and K. Ponnuraj, *J. Mol. Biol.*, 2010, **400**, 274–283.

57. F. Basolo and R. G. Pearson, *Mechanisms of Inorganic Reactions*, John Wiley & Sons, Inc., New York, 2nd edn, 1967.

58. I.-S. Park and R. P. Hausinger, *Protein Sci.*, 1993, **2**, 1034–1041.
59. E. Jabri and P. A. Karplus, *Biochemistry*, 1996, **35**, 10616–10626.
60. I.-S. Park and R. P. Hausinger, *J. Protein Chem.*, 1993, **12**, 51–56.
61. I.-S. Park and R. P. Hausinger, *Biochemistry*, 1996, **35**, 5345–5352.
62. M. J. Todd and R. P. Hausinger, *J. Biol. Chem.*, 1991, **266**, 24327–24331.
63. M. J. Todd and R. P. Hausinger, *J. Biol. Chem.*, 1991, **266**, 10260–10267.
64. P. R. Martin and R. P. Hausinger, *J. Biol. Chem.*, 1992, **267**, 20024–20027.
65. M. A. Pearson, L. Overbye Michel, R. P. Hausinger and P. A. Karplus, *Biochemistry*, 1997, **36**, 8164–8172.
66. M. A. Pearson, R. A. Schaller, L. O. Michel, P. A. Karplus and R. P. Hausinger, *Biochemistry*, 1998, **37**, 6214–6220.
67. Z. Amtul, A.-U. Rahman, R. A. Siddiqui and M. I. Choudhary, *Curr. Med. Chem.*, 2002, **9**, 1323–1348.
68. M. G. Finnegan, A. T. Kowal, M. T. Werth, P. A. Clark, D. E. Wilcox and M. K. Johnson, *J. Am. Chem. Soc.*, 1991, **113**, 4030–4032.
69. M. J. Todd and R. P. Hausinger, *J. Biol. Chem.*, 1989, **264**, 15835–15842.
70. S. Benini, S. Ciurli, W. R. Rypniewski, K. S. Wilson and S. Mangani, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1998, **54**, 409–412.
71. S. Benini, W. R. Rypniewski, K. S. Wilson, S. Ciurli and S. Mangani, *J. Biol. Inorg Chem.*, 1998, **3**, 268–273.
72. G. B. Kistiakowsky and R. Lumry, *J. Am. Chem. Soc.*, 1949, **71**, 2006–2013.
73. J. F. Ambrose, G. B. Kistiakowsky and A. G. Kridl, *J. Am. Chem. Soc.*, 1950, **72**, 317–321.
74. L. Mazzei, M. Cianci, S. Benini, L. Bertini, F. Musiani and S. Ciurli, *J. Inorg. Biochem.*, 2016, **154**, 42–49.
75. K. Kobashi, J.-I. Hase and K. Uehara, *Biochim. Biophys. Acta*, 1962, **62**, 380–383.
76. N. E. Dixon, C. Gazzola, J. J. Watters, R. Blakeley and B. Zerner, *J. Am. Chem. Soc.*, 1975, **97**, 4130–4131.
77. S. Benini, W. R. Rypniewski, K. S. Wilson, S. Miletti, S. Ciurli and S. Mangani, *J. Biol. Inorg Chem.*, 2000, **5**, 110–118.
78. G. Lethbridge and R. G. Burns, *Soil Biol. Biochem.*, 1976, **8**, 99–102.
79. C. B. Christianson, B. H. Byrnes and G. Carmona, *Fert. Res.*, 1990, **26**, 21–27.
80. G. L. Creason, M. R. Schmidt, E. A. Douglass and L. L. Hendrickson, *Soil Biol. Biochem.*, 1990, **22**, 209–211.
81. R. K. Andrews, A. Dexter, R. L. Blakeley and B. Zerner, *J. Am. Chem. Soc.*, 1986, **108**, 7124–7125.
82. G. W. McCarthy, J. M. Bremner and J. S. Lee, *Plant Soil*, 1990, **127**, 269–283.
83. W. S. Faraci, B. V. Yang, D. O'Rourke and R. W. Spencer, *Bioorg. Med. Chem.*, 1995, **3**, 605–610.
84. A. J. Pope, C. D. Toseland, B. Rushant, S. Richardson, M. McVey and J. Hills, *Dig. Dis. Sci.*, 1998, **43**, 109–119.

85. M. J. Dominguez, C. Sanmartin, M. Font, J. A. Palop, S. San Francisco, O. Urrutia, F. Houdusse and J. Garcia-Mina, *J. Agric. Food Chem.*, 2008, **56**, 3721–3731.

86. S. Vassiliou, A. Grabowiecka, P. Kosikowska, A. Yiotakis, P. Kafarski and L. Berlicki, *J. Med. Chem.*, 2008, **51**, 5736–5744.

87. S. Vassiliou, P. Kosikowska, A. Grabowiecka, A. Yiotakis, P. Kafarski and L. Berlicki, *J. Med. Chem.*, 2010, **53**, 5597–5606.

88. S. P. Howell and J. B. Sumner, *J. Biol. Chem.*, 1934, **104**, 619–626.

89. B. Krajewska and W. Zaborska, *J. Mol. Catal.*, 1999, **6B**, 75–81.

90. S. Benini, W. R. Rypniewski, K. S. Wilson, S. Ciurli and S. Mangani, *J. Biol. Inorg Chem.*, 2001, **6**, 778–790.

91. S. Benini, W. R. Rypniewski, K. S. Wilson, S. Mangani and S. Ciurli, *J. Am. Chem. Soc.*, 2004, **126**, 3714–3715.

92. B. Krajewska, W. Zaborska, M. Leszko and Z. Brzozka, *Pol. J. Chem.*, 1999, **73**, 359–366.

93. J. M. Breitenbach and R. H. Hausinger, *Biochem. J.*, 1988, **250**, 917–920.

94. S. Benini, P. Kosikowska, M. Cianci, L. Mazzei, A. G. Vara, L. Berlicki and S. Ciurli, *J. Biol. Inorg Chem.*, 2013, **18**, 391–399.

95. M. J. Todd and R. P. Hausinger, *Biochemistry*, 2000, **39**, 5389–5396.

96. B. Krajewska, W. Zaborska and M. Leszko, *J. Mol. Catal. B: Enzym.*, 2001, **14**, 101–109.

97. S. Benini, M. Cianci, L. Mazzei and S. Ciurli, *J. Biol. Inorg Chem.*, 2014, **19**, 1243–1261.

98. B. Krajewska, *J. Enzyme Inhib. Med. Chem.*, 2008, **23**, 535–542.

99. C. Follmer and C. R. Carlini, *Arch. Biochem. Biophys.*, 2005, **435**, 15–20.

100. J. R. Lambert and P. Midolo, *Aliment. Pharmacol. Ther.*, 1997, **11**(suppl. 1), 27–33.

101. G. G. Briand and N. Burford, *Chem. Rev.*, 1999, **99**, 2601–2658.

102. H. Sun, L. Zhang and K. Y. Szeto, *Met. Ions Biol. Syst.*, 2004, **41**, 333–378.

103. L. Zhang, S. B. Mulrooney, A. F. K. Leung, Y. Zeng, B. B. C. Ko, R. P. Hausinger and H. Sun, *BioMetals*, 2006, **19**, 503–511.

104. J. W. Lown, *Mol. Cell. Biochem.*, 1983, **55**, 17–40.

105. F. Geike, *Z. Pflanzenernaehr. Bodenkd.*, 1972, **132**, 256–257.

106. M. G. Peter, *Angew. Chem., Int. Ed.*, 1989, **28**, 555–570.

107. W. Zaborska, B. Krajewska, M. Kot and W. Karcz, *Bioorg. Chem.*, 2007, **35**, 233–242.

108. J. M. Bremner and L. A. Douglas, *Soil Biol. Biochem.*, 1971, **3**, 297–307.

109. L. G. Bundy and J. M. Bremner, *Soil Biol. Biochem.*, 1973, **5**, 847–853.

110. R. L. Mulvaney and J. M. Bremner, *Soil Biol. Biochem.*, 1978, **10**, 297–302.

111. L. Mazzei, M. Cianci, F. Musiani and S. Ciurli, *Dalton Trans.*, 2016, **45**, 5455–5459.

112. Z. P. Xiao, D. H. Shi, H. Q. Li, L. N. Zhang, C. Xu and H. L. Zhu, *Bioorg. Med. Chem.*, 2007, **15**, 3703–3710.

113. Z. P. Xiao, T. W. Ma, W. C. Fu, X. C. Peng, A. H. Zhang and H. L. Zhu, *Eur. J. Med. Chem.*, 2010, **45**, 5064–5070.

114. Z. P. Xiao, Z. Y. Peng, J. J. Dong, J. He, H. Ouyang, Y. T. Feng, C. L. Lu, W. Q. Lin, J. X. Wang, Y. P. Xiang and H. L. Zhu, *Eur. J. Med. Chem.*, 2013, **63**, 685–695.

115. S. Matsubara, H. Shibata, F. Ishikawa, T. Yokokura, M. Takahashi, T. Sugimura and K. Wakabayashi, *Biochem. Biophys. Res. Commun.*, 2003, **310**, 715–719.

116. M. Kot and W. Zaborska, *J. Enzyme Inhib. Med. Chem.*, 2003, **18**, 413–417.

117. J. Yang, M. A. Cohen Stuart and M. Kamperman, *Chem. Soc. Rev.*, 2014, **43**, 8271–8298.

118. G. Li, H. Zhang, F. Sader, N. Vadhavkar and D. Njus, *Biochemistry*, 2007, **46**, 6978–6983.

119. L. Mazzei, M. Cianci, F. Musiani, G. Lente, M. Palombo and S. Ciurli, *J. Inorg. Biochem.*, 2017, **166**, 182–189.

120. S. Ciurli, in *Nickel and its Surprising Impact in Nature*, ed. A. Sigel, H. Sigel and R. K. O. Sigel, John Wiley & Sons, Ltd., Chichester, UK, 2007, vol. 10, pp. 241–278.

121. F. Musiani, E. Arnofi, R. Casadio and S. Ciurli, *J. Biol. Inorg Chem.*, 2001, **6**, 300–314.

122. A. M. Barrios and S. J. Lippard, *J. Am. Chem. Soc.*, 1999, **121**, 11751–11757.

123. D. Suarez, N. Diaz and K. M. Merz, *J. Am. Chem. Soc.*, 2003, **125**, 15324–15337.

124. M. A. Pearson, I. S. Park, R. A. Schaller, L. O. Michel, P. A. Karplus and R. P. Hausinger, *Biochemistry*, 2000, **39**, 8575–8584.

125. A. E. Martell and R. M. Smith, *Critical Stability Constants*, Plenum Press, New York, 1974.

126. M. Laitaoja, J. Valjakka and J. Janis, *Inorg. Chem.*, 2013, **52**, 10983–10991.

127. N. E. Dixon, C. Gazzola, C. J. Asher, D. W. Lee, R. L. Blakeley and B. Zerner, *Can. J. Biochem.*, 1980, **58**, 474–480.

128. G. J. King and B. Zerner, *Inorg. Chim. Acta*, 1997, **255**, 381–388.

129. K. Yamaguchi, N. J. Cosper, C. Stalhandske, R. A. Scott, M. A. Pearson, P. A. Karplus and R. P. Hausinger, *J. Biol. Inorg Chem.*, 1999, **4**, 468–477.

130. J. Stoof, S. Breijer, R. G. Pot, D. van der Neut, E. J. Kuipers, J. G. Kusters and A. H. van Vliet, *Environ. Microbiol.*, 2008, **10**, 2586–2597.

131. E. L. Carter, D. E. Tronrud, S. R. Taber, P. A. Karplus and R. P. Hausinger, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 13095–13099.

132. A. B. Becker-Ritt, A. H. Martinelli, S. Mitidieri, V. Feder, G. E. Wassermann, L. Santi, M. H. Vainstein, J. T. Oliveira, L. M. Fiuza, G. Pasquali and C. R. Carlini, *Toxicon*, 2007, **50**, 971–983.

133. A. B. Becker-Ritt and C. R. Carlini, *Biopolymers*, 2012, **98**, 367–384.

134. F. Staniscuaski, V. Te Brugge, C. R. Carlini and I. Orchard, *J. Insect Physiol.*, 2009, **55**, 255–263.

135. F. Staniscuaski, V. Te Brugge, C. R. Carlini and I. Orchard, *J. Insect Physiol.*, 2010, **56**, 1078–1086.

136. M. S. Defferrari, R. da Silva, I. Orchard and C. R. Carlini, *Toxicon*, 2014, **82**, 18–25.

137. M. S. Defferrari, D. H. Lee, C. L. Fernandes, I. Orchard and C. R. Carlini, *Biochim. Biophys. Acta*, 2014, **1840**, 396–405.

138. C. R. Carlini, J. A. Guimaraes and J. M. Ribeiro, *Br. J. Pharmacol.*, 1985, **84**, 551–560.

139. C. Barja-Fidalgo, J. A. Guimaraes and C. R. Carlini, *Endocrinology*, 1991, **128**, 675–679.

140. C. Barja-Fidalgo, J. A. Guimaraes and C. R. Carlini, *Toxicon*, 1991, **29**, 453–459.

141. F. A. Ghazaleh, I. M. Francischetti, M. E. Gombarovits and C. R. Carlini, *Arch. Biochem. Biophys.*, 1997, **339**, 362–367.

142. C. R. Carlini and J. A. Guimaraes, *Toxicon*, 1981, **19**, 667–675.

143. C. Follmer, G. B. S. Barcellos, R. B. Zingali, O. L. T. Machado, E. W. Alves, C. Barja-Fidalgo, J. A. Guimaraes and C. R. Carlini, *Biochem. J.*, 2001, **360**, 217–224.

144. F. Mulinari, F. Staniscuaski, L. R. Bertholdo-Vargas, M. Postal, O. B. Oliveira-Neto, D. J. Rigden, M. F. Grossi-de-Sa and C. R. Carlini, *Peptides*, 2007, **28**, 2042–2050.

145. F. C. Lopes, O. Dobrovolska, R. Real-Guerra, V. Broll, B. Zambelli, F. Musiani, V. N. Uversky, C. R. Carlini and S. Ciurli, *FEBS J.*, 2015, **282**, 1043–1064.

146. H. L. Ehrlich, *Earth-Sci. Rev.*, 1998, **45**, 45–60.

147. S. Castanier, G. Le Métayer-Levrel and J.-P. Perthuisot, *Sediment. Geol.*, 1999, **126**, 9–23.

148. S. Douglas and T. J. Beveridge, *FEMS Microbiol. Ecol.*, 1998, **26**, 79–88.

149. J. T. DeJong, B. M. Mortensen, B. C. Martinez and D. C. Nelson, *Ecol. Eng.*, 2010, **36**, 197–210.

150. J. T. DeJong, K. Soga, E. Kavazanjian, S. Burns, L. A. van Paassen, A. Al Qabany, A. Aydilek, S. S. Bang, M. Burbank, L. F. Caslake, C. Y. Chen, X. Cheng, J. Chu, S. Ciurli, A. Esnault-Filet, S. Fauriel, N. Hamdan, T. Hata, Y. Inagaki, S. Jefferis, M. Kuo, L. Laloui, J. Larrahondo, D. A. C. Manning, B. Martinez, B. M. Montoya, D. C. Nelson, A. Palomino, P. Renforth, J. C. Santamarina, E. A. Seagren, B. Tanyu, M. Tsesarsky and T. Weaver, *Géotechnique*, 2013, **63**, 287–301.

151. M. A. Farrugia, L. Macomber and R. P. Hausinger, *J. Biol. Chem.*, 2013, **288**, 13178–13185.

152. Z. Chang, J. Kuchar and R. P. Hausinger, *J. Biol. Chem.*, 2004, **279**, 15305–15313.

153. S. R. Heimer and H. L. T. Mobley, *J. Bacteriol.*, 2001, **183**, 1423–1433.

154. I. S. Park, M. B. Carr and R. P. Hausinger, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 3233–3237.

155. F. Musiani, M. Bellucci and S. Ciurli, *J. Chem. Inf. Model.*, 2011, **51**, 1513–1520.

156. M. Salomone-Stagni, B. Zambelli, F. Musiani and S. Ciurli, *Proteins*, 2007, **68**, 749–761.

157. E. L. Carter, J. L. Boer, M. A. Farrugia, N. Flugga, C. L. Towns and R. P. Hausinger, *Biochemistry*, 2011, **50**, 9296–9308.

158. B. Zambelli, A. Berardi, V. Martin-Diaconescu, L. Mazzei, F. Musiani, M. J. Maroney and S. Ciurli, *J. Biol. Inorg Chem.*, 2014, **19**, 319–334.
159. A. Soriano and R. P. Hausinger, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 11140–11144.
160. B. Zambelli, M. Stola, F. Musiani, K. De Vriendt, B. Samyn, B. Devreese, J. Van Beeumen, P. Turano, A. Dikiy, D. A. Bryant and S. Ciurli, *J. Biol. Chem.*, 2005, **280**, 4684–4695.
161. F. Musiani, E. Ippoliti, C. Micheletti, P. Carloni and S. Ciurli, *Biochemistry*, 2013, **52**, 2949–2954.
162. B. Zambelli, P. Turano, F. Musiani, P. Neyroz and S. Ciurli, *Proteins*, 2009, **74**, 222–239.
163. V. Martin-Diaconescu, M. Bellucci, F. Musiani, S. Ciurli and M. J. Maroney, *J. Biol. Inorg Chem.*, 2012, **17**, 353–361.
164. J. C. Rain, L. Selig, H. De Reuse, V. Battaglia, C. Reverdy, S. Simon, G. Lenzen, F. Petel, J. Wojcik, V. Schachter, Y. Chemama, A. Labigne and P. Legrain, *Nature*, 2001, **409**, 211–215.
165. P. Voland, D. L. Weeks, E. A. Marcus, C. Prinz, G. Sachs and D. Scott, *Am. J. Physiol.: Gastrointest. Liver Physiol.*, 2003, **284**, G96–G106.
166. M. Bellucci, B. Zambelli, F. Musiani, P. Turano and S. Ciurli, *Biochem. J.*, 2009, **422**, 91–100.
167. A. Merloni, O. Dobrovolska, B. Zambelli, F. Agostini, M. Bazzani, F. Musiani and S. Ciurli, *Biochim. Biophys. Acta*, 2014, **1844**, 1662–1674.
168. R. Lam, V. Romanov, K. Johns, K. P. Battaile, J. Wu-Brown, J. L. Guthrie, R. P. Hausinger, E. F. Pai and N. Y. Chirgadze, *Proteins*, 2010, **78**, 2839–2848.
169. Y. H. Fong, H. C. Wong, C. P. Chuck, Y. W. Chen, H. Sun and K. B. Wong, *J. Biol. Chem.*, 2011, **286**, 43241–43249.
170. Y. H. Fong, H. C. Wong, M. H. Yuen, P. H. Lau, Y. W. Chen and K. B. Wong, *PLoS Biol.*, 2013, **11**, e1001678.
171. H. Remaut, N. Safarov, S. Ciurli and J. Van Beeumen, *J. Biol. Chem.*, 2001, **276**, 49365–49370.
172. B. Zambelli, K. Banaszak, A. Merloni, A. Kiliszek, W. Rypniewski and S. Ciurli, *J. Biol. Inorg Chem.*, 2013, **18**, 1005–1017.
173. H. K. Song, S. B. Mulrooney, R. Huber and R. P. Hausinger, *J. Biol. Chem.*, 2001, **276**, 49359–49364.
174. R. Shi, C. Munger, A. Asinas, S. L. Benoit, E. Miller, A. Matte, R. J. Maier and M. Cygler, *Biochemistry*, 2010, **49**, 7080–7088.
175. K. Banaszak, V. Martin-Diaconescu, M. Bellucci, B. Zambelli, W. Rypniewski, M. J. Maroney and S. Ciurli, *Biochem. J.*, 2012, **441**, 1017–1026.
176. F. Musiani, B. Zambelli, M. Stola and S. Ciurli, *J. Inorg. Biochem.*, 2004, **98**, 803–813.
177. M. Stola, F. Musiani, S. Mangani, P. Turano, N. Safarov, B. Zambelli and S. Ciurli, *Biochemistry*, 2006, **45**, 6495–6509.
178. X. Yang, H. Li, T.-P. Lai and H. Sun, *J. Biol. Chem.*, 2015, **290**, 12474–12485.
179. M. A. Farrugia, B. Wang, M. Feig and R. P. Hausinger, *Biochemistry*, 2015, **54**, 6392–6401.

180. P. W. O'Toole, W. J. Snelling, C. Canchaya, B. M. Forde, K. R. Hardie, C. Josenhans, R. L. Graham, G. McMullan, J. Parkhill, E. Belda and S. D. Bentley, *BMC Genomics*, 2010, **11**, 1–24.

181. J. K. Kim, S. B. Mulrooney and R. P. Hausinger, *J. Bacteriol.*, 2005, **187**, 7150–7154.

182. A. Balasubramanian, V. Durairajpandian, S. Elumalai, N. Mathivanan, A. K. Munirajan and K. Ponnuraj, *Int. J. Biol. Macromol.*, 2013, **58**, 301–309.

ANNEX 2

Reprinted from:

Journal: Journal of Biological Inorganic Chemistry

Title: Fluoride inhibition of *Sporosarcina pasteurii* urease: structure and thermodynamics

Volume: 19

Year: 2014

Pages: 1243 - 1261

Authors: S. Benini, M. Cianci, L. Mazzei and S. Ciurli

ORIGINAL PAPER

Fluoride inhibition of *Sporosarcina pasteurii* urease: structure and thermodynamics

Stefano Benini · Michele Cianci · Luca Mazzei ·
Stefano Ciurli

Abstract Urease is a nickel-dependent enzyme and a virulence factor for ureolytic bacterial human pathogens, but it is also necessary to convert urea, the most worldwide used fertilizer, into forms of nitrogen that can be taken up by crop plants. A strategy to control the activity of urease for medical and agricultural applications is to use enzyme inhibitors. Fluoride is a known urease inhibitor, but the structural basis of its mode of inhibition is still undetermined. Here, kinetic studies on the fluoride-induced inhibition of urease from *Sporosarcina pasteurii*, a widespread and highly ureolytic soil bacterium, were performed using isothermal titration calorimetry and revealed a mixed competitive and uncompetitive mechanism. The pH dependence of the inhibition constants, investigated in the 6.5–8.0 range, reveals a predominant uncompetitive mechanism that increases by increasing the pH, and a lesser competitive inhibition that increases by lowering the pH. Ten crystal structures of the enzyme were independently determined using five crystals of the native form and five crystals of the protein crystallized in the presence of fluoride. The analysis of these structures revealed the presence of two fluoride anions coordinated to the Ni(II) ions in the active site, in terminal and bridging positions. The present study consistently supports an interaction of fluoride with the nickel centers in the urease active site in which one fluoride competitively binds to the Ni(II) ion proposed to coordinate urea in the initial step of the catalytic mechanism, while another fluoride uncompetitively substitutes the Ni(II)-bridging hydroxide, blocking its nucleophilic attack on urea.

Keywords X-ray crystallography · Urease · Fluoride · Nickel · *Sporosarcina pasteurii* · Enzyme inhibition · Calorimetry

An Interactive 3D Complement page in Proteopedia is available at: http://proteopedia.org/wiki/index.php/Journal:JBIC:26.

S. Benini (✉)
Faculty of Science and Technology, Free University of Bolzano,
Piazza Università 5, 39100 Bolzano, Italy
e-mail: stefano.benini@unibz.it

M. Cianci
European Molecular Biology Laboratory c/o DESY,
Notkestrasse 85, 22603 Hamburg, Germany

L. Mazzei · S. Ciurli (✉)
Laboratory of Bioinorganic Chemistry, Department of Pharmacy
and Biotechnology, University of Bologna, Via Fanin 40,
40127 Bologna, Italy
e-mail: stefano.ciurli@unibo.it

Introduction

Urease (E.C. 3.5.1.5) is a nickel-dependent non-redox enzyme [1] that catalyzes the hydrolysis of urea (Scheme 1) in the last step of organic nitrogen mineralization in bacteria, fungi, plants, algae and invertebrates [2–5] by increasing the rate of urea hydrolysis 10^{15} times as compared to the non-catalyzed hydrolysis reaction [6].

The high urease activity triggers an overall increase in pH of the surrounding milieu, which in turn causes negative consequences on human and animal health as well as in plant crop production [2–5, 7–10]. Therefore, efficient urease inhibitors are needed for both medical and agricultural applications. These inhibitors could target either the urease activation process, which involves the interaction of four accessory proteins named UreD, UreF, UreG and UreE with the urease apoenzyme, resulting in the insertion of the two Ni(II) ions in the active site [5, 11], or the

🖄 Springer

Scheme 1 Urease catalysed hydrolysis
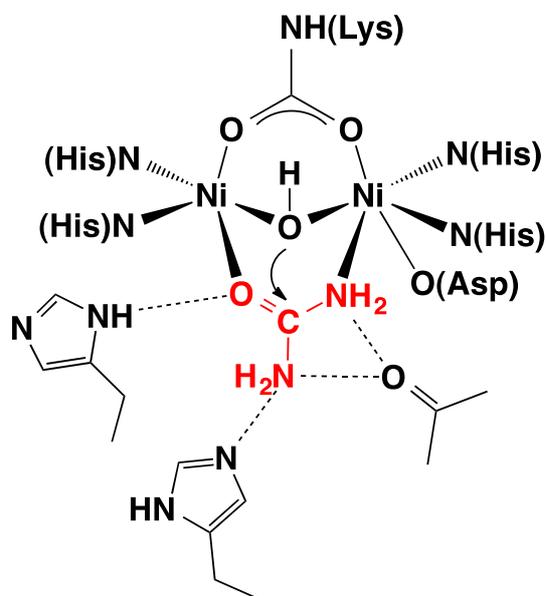
Scheme 2 Nucleophilic attack to urea carbon by the bridging hydroxide upon urea binding to the Ni ions

activated enzyme itself. In the latter case, knowledge of the structure of the nickel-containing catalytic cavity is essential. Several structural studies on ureases from different biological sources [12–16] have revealed that the immediate environment around the two Ni(II) ions at the active site is conserved, as to induce a common mechanism of catalysis whose key step is the nucleophilic attack of the nickel-bridging hydroxide on the urea molecule bound to the bimetallic nickel cluster via O and N atoms (Scheme 2) [1, 5, 7, 9, 10, 13]. In this mechanism, the reaction is accompanied by the formation of an extended network of second shell hydrogen bonds that stabilizes the binding of the substrate, and by the movement of a flexible flap that changes the active site channel from an open to a closed conformation.

Several classes of inhibitors have been proposed and tested [4, 17]. Urea structural analogs, such as hydroxyurea, formamide, thiourea and alkyl-substituted urea are competitive inhibitors, and the same behavior is observed for thiols, phosphate, boric and boronic acids, as well as Bi(III) complexes [4, 17]. Hydroxamates, heavy metal ions, as well as amide and ester derivatives of phosphoric and thiophosphoric acids compose another group of competitive and slow-binding inhibitors, some of which have found applications in medicine and agriculture because of their

high efficacy; however, they exhibit either toxicity or low stability and are therefore not optimal for these applications [4, 17]. Most recently, the synthesis and study of a novel class of inhibitors based on phosphinic and thiophosphinic acid skeletons, which contain a hydrolytically stable C–P bond, have been reported and represent promising leads for further developments [18–20]. The structural basis for the inhibitii properties of some of these classes of inhibitors has been established by a series of crystal structures of ureases from different sources bound to β-mercaptoethanol [21], acetohydroxamic acid [14, 22, 23], phosphate [15, 24], boric acid [25], and diamidophosphate [13]. Recently, the structure of a urease bound to citrate revealed the potential for carboxylates to act as competitive enzyme inhibitors [26].

Fluoride is a good ligand for metal ions and has been reported to act as a mixed competitive/uncompetitive [27], pseudo-uncompetitive slow-binding [28], competitive [29], and competitive slow-binding [30] urease inhibitor in a number of kinetic studies. Here, we discuss crystallographic evidences for the structure of the complex formed between fluoride and urease from *S. pasteurii* (SPU) as compared to the structure of the same enzyme in the native state. This analysis, concomitantly with a re-evaluation of the behavior of fluoride as inhibitor of SPU using a calorimetry-based enzymatic assay, allowed us to accurately identify the active site Ni(II) ligands as solvent molecules or fluoride ions, and therefore to draw a rationale for the mechanism of inhibition. This study supports the view that fluoride inhibits urease using a mixed competitive and uncompetitive mechanism. This action is exerted by binding to the active site Ni(II) ion that is thought to be initially involved in the binding of urea, as well as by replacing the nickel-bridging hydroxide that appears to act as the nucleophile in the urea hydrolysis reaction [5, 7, 9, 10, 13].

Materials and methods

Sporosarcina pasteurii (previously known as *Bacillus pasteurii*) urease (SPU) was purified using a previously reported protocol [26].

Kinetic studies

Calorimetric experiments were carried out with a high-sensitivity VP-ITC (ITC: isothermal titration calorimetry) micro-calorimeter (MicroCal LLC, Northampton, MA, USA), using a previously described enzymatic assay [31, 32]. The calorimeter is made of a reference cell filled with deionized water and a sample cell where the reaction occurs. Reference and sample cells are maintained at the

same temperature. The method monitors the enzymatic conversion of substrate by probing the heat (Q, cal) generated over time, defined as thermal power TP (cal s^{-1}), necessary to maintain the reaction cell at the same constant temperature of the reference cell:

$$TP = \frac{dQ}{dt} \tag{1}$$

The heat (Q) is proportional to the moles of substrate urea converted to products (n) and to the molar enthalpy of the reaction (ΔH, cal mol^{-1}):

$$Q = n \times \Delta H \tag{2}$$

The number of moles can be calculated as the product of the molar concentration of the converted substrate ([urea], mol L^{-1}) times the total volume (V, L) of the solution, and therefore the overall heat Q can be calculated:

$$Q = [\text{urea}] \times V \times \Delta H \tag{3}$$

The reaction rate, defined as the change in substrate concentration over time can thus be related to the thermal power:

$$v = -\frac{d[\text{urea}]}{dt} = -\frac{1}{V \times \Delta H} \times \frac{dQ}{dt} = -\frac{1}{V \times \Delta H} \times TP \tag{4}$$

According to the above equation, the reaction rate can be determined by first establishing the reaction molar enthalpy ΔH in an experiment defined as M1, and then by monitoring the thermal power upon multiple substrate additions, with an experiment defined as M2.

For both M1 and M2, the SPU samples were eluted through a Superdex 200 10/300 GL (GE Healthcare) size-exclusion chromatography column immediately before the measurements, using a 50 mM HEPES pH 7.0 buffer, containing 50 mM Na$_2$SO$_3$ and 150 mM NaCl, as eluent. The protein samples were then diluted with 50 mM HEPES (at the pH value used in the experiment, 6.5, 7.0 or 8.0) to the concentrations used for M1 and M2 (indicated below), degassed, and loaded into the ITC sample cell ($V = 1.4093$ mL). The ITC stirring injection syringe was filled with urea dissolved in the same buffer to the concentrations used for M1 and M2 (see below). The reference cell was filled with deionized water, and the temperature of the two cells was set and stabilized at 298 K. Stirring speed was 300 rpm, and thermal power was monitored every 2 s using high instrumental feedback.

For experiment M1, 30 nM SPU and 20 mM urea were used in cell and in syringe, respectively, and a single injection of 5 µL urea was carried out, giving a final substrate concentration of 0.07 mM in the sample cell. After the baseline returned to the original level, indicating that the consumption of the substrate was complete, a second injection was performed. Numerical integration of the area under the two single peaks was carried out, and the average value thus obtained yielded the molar enthalpy for urea hydrolysis in the utilized experimental conditions (buffer composition, ionic strength, temperature and pH).

For experiment M2, 45–105 pM SPU and 0.4–1.0 M urea were used in the cell and in the syringe, respectively. After 15 min of enzyme pre-incubation at 25 °C in 50 mM HEPES pH 6.5, 7.0 or 8.0 with or without fluoride, depending on the experiment (see below), successive injections of 5–7 µL urea were carried out every 2–3 min, a time necessary to allow the thermal equilibrium to return to a steady-state level after each injection and, at the same time, maintain pseudo-first-order reaction conditions. The thermal power obtained from the baseline shift was averaged using the last 15 s prior to the subsequent injection to obtain an accurate measurement. The calculated thermal power for each injection was converted to reaction rate using Eq. (4) and corrected for enzyme concentration. Experiments were conducted in the same conditions in the absence and in the presence of 100–800 µM NaF, dissolved in both the enzyme and the substrate solutions to keep the concentration constant. Control experiments were carried out by injecting the urea solution into the buffer alone. Data were processed with the Origin package provided by calorimeter manufacturer. The reaction rates, in the absence and in the presence of fluoride, were obtained from the change in thermal power calculated as the difference between the original baseline and the new baseline following each injection, using Eq. (4) and the value of ΔH determined in M1.

The data analysis was based on the following definitions of types of enzyme inhibition [33]:

1. the inhibitor binds to the enzyme independently of the presence of substrate, in a reversible equilibrium of the type E + I $\rightleftarrows$ EI; for this competitive inhibition the dissociation constant is expressed as K_{ic}; in this case, the rate of enzymatic reaction is given by the Eq. (5), in which $V_{max} = k_{cat}[\text{urease}]$ and K_m are the maximum rate and the Michaelis constant, respectively, while $[I]$ is the molar concentration of the inhibitor:

$$v = \frac{V_{max} \times [\text{urea}]}{[\text{urea}] + K_m \times \left(1 + \frac{[I]}{K_{ic}}\right)} \tag{5}$$

2. the inhibitor binds to the enzyme only in the presence of substrate in an uncompetitive inhibition equilibrium of the type ES + I $\rightleftarrows$ ESI, for which the dissociation constant is expressed as K_{iuc}; in this case, the rate of enzymatic reaction is given by:

$$v = \frac{V_{\max} \times [\text{urea}]}{K_{\text{m}} + [\text{urea}] \times \left(1 + \frac{[I]}{K_{\text{iuc}}}\right)} \qquad (6)$$

3. the inhibitor binds to the enzyme both in the presence and in the absence of substrate, so that both equilibria above are operating, in the so-called non-competitive, or mixed, inhibition; in this case, the rate of enzymatic reaction is given by:

$$v = \frac{V_{\max} \times [\text{urea}]}{K_{\text{m}} \times \left(1 + \frac{[I]}{K_{\text{ic}}}\right) + [\text{urea}] \times \left(1 + \frac{[I]}{K_{\text{iuc}}}\right)} \qquad (7)$$

The kinetic parameters were obtained from a fit of the experimental reaction rates obtained by calorimetry to the general inhibition Eq. (7), using non-linear regression analysis implemented in MacCurveFit (v. 1.5.4 Kevin Raner software).

Crystallographic studies

Crystallization trials were performed at 293 K using the hanging-drop method and 1 μL of a 11 mg/mL SPU solution in 20 mM Na-HEPES, pH = 7.0, containing 50 mM Na$_2$SO$_3$, which was diluted with 1 μL of the precipitant solution. The drop was equilibrated by vapor diffusion against 1 mL of the precipitant solution using a Hampton Research 24-well Linbro plate. Protein crystals of native SPU appeared within 1–3 weeks and grew to a size of $0.3 \times 0.3 \times 0.6$ mm^3 when they were equilibrated against a solution containing 1.6–1.8 M ammonium sulfate in a 100 mM sodium citrate buffer and in 50 mM Na$_2$SO$_3$ (final pH $\sim$6.5 [26]). Protein crystals of fluoride-inhibited SPU were obtained using identical conditions except for a precipitant solution that additionally contained 100 mM sodium fluoride.

Crystals of native and fluoride-bound SPU were scooped up using cryoloops, transferred to a cryoprotectant solution containing 20 % ethylene glycol, 2.4 M ammonium sulfate, 100 mM sodium citrate and 50 mM Na$_2$SO$_3$ (100 mM sodium fluoride was additionally present in the case of the crystallization of the fluoride complex) and then flash cooled and stored in liquid nitrogen. Diffraction data were collected at 100 K using synchrotron radiation at the EMBL P13 beamline of the Petra III storage ring, c/o DESY, Hamburg (Germany). The wavelength was set to 0.968 Å using a Si(111) crystal monochromator (FMB-Oxford). The beamline was equipped with a Rayonix HE225 CCD detector and a MD2 goniometer (Maatel-EMBL) with an horizontal spindle axis. Reflection data were collected in two sweeps to accurately record low- and high-resolution data without overloading the detector. Each sweep consisted of 360 images with 0.2° oscillation. The

data were processed using XDS [34] and SCALA [35]. The complete diffraction data statistics of five crystals of native urease and five crystals of urease obtained in the presence of fluoride are given in Tables 1 and 2. All the crystals were highly isomorphic with space group $P6_322$; the number of molecules per unit cell (Z-number) was 12, with a resulting solvent content of 55 % and a Matthews coefficient of 2.73 Å^3/Da.

In all cases, the structure solution and refinement followed exactly the same protocol. The model of SPU in complex with citrate (PDB code 4AC7), devoid of water and ligands, was used as a starting model for the rigid body refinement of the single subunits, conducted using Refmac [36, 37]. The model building and the water or ligand addition/inspection were conducted using Coot [38, 39]. The structure was isotropically refined, including the hydrogen atoms in the riding positions. The final refinement statistics are reported in Tables 1 and 2. The highest overall quality structures of native and fluoride-inhibited SPU were deposited in the RCSB Protein Data Bank with accession codes 4CEU and 4CEX, respectively, while all remaining crystallographic data are available upon request. Crystallographic figures were created using PyMOL (The PyMOL Molecular Graphics System, Schroedinger, LLC.) or Chimera [40].

Results

Calorimetric analysis

The inhibition of *S. pasteurii* urease by fluoride was investigated using isothermal titration calorimetry (ITC), a method that is becoming a major tool to aid enzyme inhibitor screening and design [31, 32, 41]. The molar reaction enthalpy was determined using the so-called M1 experiment, in which a single injection of diluted substrate solution into a concentrated solution of the enzyme caused a decrease of the instrumental thermal power necessary to maintain the temperature constant, indicating an exothermic reaction. Complete consumption of the substrate occurred within ca. 25 min (Fig. 1a). The integration of the curve at pH 7.0 yielded $\Delta H = -10.0$ kcal mol^{-1}. A second injection of substrate into the reaction cell (now containing ammonium ions and bicarbonate, the products of the reaction) provided a curve with identical shape and area, confirming the value of ΔH and showing negligible inhibition by products (Fig. 1a). The different value of ΔH measured in the present study at pH 7.0, as compared to that obtained previously in the case of jack bean urease (JBU) ($\Delta H = -2.8$ kcal mol^{-1}) [29], is ascribed to the different buffers used in the two calorimetric experiments (HEPES vs. Tris), which have different ionization enthalpies [ΔH_{ion} (HEPES) $= -4.88$ kcal mol^{-1}); ΔH_{ion}

Table 1 X-ray diffraction data collection and refinement statistics for native (PDB code 4CEU) SPU

	Native 4CEU	Crystal native #2	Crystal native #3	Crystal native #4	Crystal native #5
Data collection					
Wavelength (Å)	0.968	1.126	1.126	1.126	1.126
Space group	$P6_322$	$P6_322$	$P6_322$	$P6_322$	$P6_322$
$a = b$ (Å)	131.12	131.18	131.66	131.17	131.85
c (Å)	188.84	189.38	189.25	189.22	189.50
Resolution (Å)[a]	97.51–1.58 (1.62–1.58)	113.6–1.47 (1.55–1.47)	114.0–1.54 (1.62–1.54)	189.22–1.63 (1.72–1.63)	189.5–1.47 (1.55–1.47)
Number of measured reflections[a]	1,277,435 (165,497)	3,858,287 (353,561)	1,190,816 (134,547)	1,153,350 (127,802)	2,390,319 (189,425)
Number of unique reflections[a]	130,522 (18,780)	161,048 (21,847)	140,347 (19,090)	117,968 (16,455)	162,414 (22,558)
R_{merge}[a,b]	0.091 (0.524)	0.106 (0.467)	0.109 (0.657)	0.094 (0.897)	0.090 (0.955)
R_{pim}[a,c]	0.030 (0.187)	0.021 (0.115)	0.039 (0.4)	0.031 (0.333)	0.022 (0.340)
$I/\sigma(I)$[a]	23.0 (4.5)	24.3 6.9	13.3 (2.0)	18.2 (2.4)	23.0 (2.6)
Multiplicity[a]	9.8 (8.8)	24 (16)	8.5 (7.0)	9.8 (7.8)	14.7 (8.4)
Completeness (%)[a]	100.0 (100.0)	98.6 (92.9)	98.9 (93.5)	99.5 (96.5)	99.4 (95.8)
B-factor from Wilson plot (Å²)	13.5	12.9	15.6	17.0	14.5
Refinement					
R, R_{free}[d] (%)	13.51, 15.49	13.23, 14.55	16.64, 19.16	14.07, 16.65	14.58, 16.90
Cruickshank's DPI[e] based on R (Å)	0.059	0.047	0.068	0.068	0.052
Cruickshank's DPI[e] based on R_{free} (Å)	0.059	0.047	0.070	0.070	0.054
ESU based on maximum likelihood (Å)	0.037	0.028	0.052	0.047	0.039
RMSD[f] of bond lengths (Å)	0.011	0.010	0.014	0.014	0.012
RMSD[f] of bond angles (°)	1.59	1.50	1.68	1.67	1.59
Average B-factor for protein atoms (Å²)	14.3	12.9	17.1	17.2	16.47
Average B-factor for water molecules (Å²)	27.7	26.5	28.6	28.7	27.7
Average B-factor for Ni atoms (Å²)	12.25, 11.35	9.93, 8.86	15.52, 14.15	14.83, 13.36	14.66, 13.30
Average B-factor for F_B (Å²)	–	–	–	–	–
Average B-factor for F_T (Å²)	–	–	–	–	–
Number of protein atoms	6,094	6,151	6,094	6,118	6,117
Number of Ni atoms	2	2	2	2	2
Number of water molecules	859	886	847	856	844
Number of fluoride ions	0	0	0	0	0
Number of sulfate molecules	6	6	5	6	5
Ramachandran most-favored region (%)	96.8	96.3	95.5	95.7	95.9
Ramachandran additional allowed region (%)	2.9	2.9	4.5	4.3	4.1

(Tris = −11.34 kcal mol⁻¹)]. Therefore, an intrinsic urea hydrolysis reaction enthalpy of −14.9 and −14.1 kcal mol⁻¹ can be derived from these two studies on SPU and JBU, respectively, consistently with the values of −14.5 determined for urease from *Helicobacter pylori* at pH 7.8 [31] and −14.7 kcal mol⁻¹ more recently reported for JBU at pH 7.0 [32].

The M2 experiment was then carried out using a diluted enzyme solution in the measuring cell and performing multiple injections of a concentrated substrate solution to maintain the substrate concentration essentially constant within the measuring time. This experiment revealed an initial increase in thermal power due to the heat of substrate dilution, followed by a decrease

Table 2 X-ray diffraction data collection and refinement statistics for fluoride-bound (PDB code 4CEX) SPU

	Fluoride 4CEX	Crystal fluoride #2	Crystal fluoride #3	Crystal fluoride #4	Crystal fluoride #5
Data collection					
Wavelength (Å)	0.968	0.968	0.968	0.968	1.127
Space group	$P6_322$	$P6_322$	$P6_322$	$P6_322$	$P6_322$
$a = b$ (Å)	131.33	131.34	131.08	131.31	131.12
c (Å)	188.88	188.69	188.39	188.90	188.64
Resolution (Å)[a]	113.74–1.59 (1.63–1.59)	113.71–1.62 (1.71–1.62)	188.47–1.72 (1.81–1.72)	113.72–1.68 (1.77–1.68)	113.6–1.63 (1.63–1.72)
Number of measured reflections[a]	1,852,002 (163,038)	1,815,217 (154,629)	1,414,007 (129,367)	962,171 (138,960)	1,165,307 (127,382)
Number of unique reflections[a]	128,776 (18,517)	121,776 (17,531)	101,225 (14,358)	109,143 (15,684)	116,761 (15,890)
$R_{\text{merge}}^{\text{a,b}}$	0.100 (0.523)	0.115 (0.492)	0.149 (0.51)	0.073 (0.47)	0.115 (0.537)
$R_{\text{pim}}^{\text{a,c}}$	0.026 (0.184)	0.030 (0.176)	0.039 (0.158)	0.026 (0.167)	0.037 (0.193)
I/σ (I)[a]	19.7 (4.6)	17.9 (4.5)	20 (4.0)	11 (5.2)	14 (3.6)
Multiplicity[a]	14.4 (8.8)	14.9 (8.8)	14.0 (9.0)	8.8 (8.9)	10 (8.0)
Completeness (%)[a]	100.0 (100.0)	100 (100)	99.5 (98.2)	99.8 (99.8)	98.4 (93)
B-factor from Wilson plot (Å^2)	12.3	15.1	15.5	14.83	13.15
Refinement					
R, $R_{\text{free}}^{\text{d}}$ (%)	12.97, 15.02	14.34, 16.21	13.63, 16.17	13.25, 15.24	12.67, 14.83
Cruickshank's DPI[e] based on R (Å)	0.057	0.066	0.077	0.069	0.061
Cruickshank's DPI[e] based on R_{free} (Å)	0.058	0.066	0.077	0.068	0.062
ESU based on maximum likelihood (Å)	0.033	0.044	0.051	0.043	0.037
RMSD[f] of bond lengths (Å)	0.011	0.014	0.015	0.013	0.012
RMSD[f] of bond angles (°)	1.53	1.68	1.72	1.62	1.54
Average B-factor for protein atoms (Å^2)	12.9	16.8	15.6	14.48	13.62
Average B-factor for water molecules (Å^2)	27.5	28.6	28.1	28.4	28.33
Average B-factor for Ni atoms (Å^2)	8.75, 9.67	12.96, 12.19	12.61, 11.52	11.89, 10.78	10.64, 9.82
Average B-factor for F_{B} (Å^2)	8.28	12.67	11.42	11.70	11.01
Average B-factor for F_{T} (Å^2)	13.6	18.48	14.15	14.81	16.54
Number of protein atoms	6,104	6,108	6,114	6,084	6,083
Number of Ni atoms	2	2	2	2	2
Number of water molecules	826	761	807	864	835
Number of fluoride ions	2	2	2	2	2
Number of sulfate molecules	3	3	4	2	3
Ramachandran most-favored region (%)	96.4	95.2	95.3	95.6	96.0
Ramachandran additional allowed region (%)	3.2	4.8	4.7	4.4	4.0

[a] The values in parentheses are for the highest-resolution shell

[b] $R_{\text{merge}} = \Sigma_{\text{hkl}} \Sigma_i |I_i(\text{hkl}) - \langle I(\text{hkl}) \rangle| / \Sigma_{\text{hkl}} \Sigma_i |I_i(\text{hkl})|$

[c] $R_{\text{pim}} = \Sigma_{\text{hkl}} (1/N - 1)^{1/2} \Sigma_i |I_i(\text{hkl}) - \langle I(\text{hkl}) \rangle| \Big/ \Sigma_{\text{hkl}} \Sigma_i |I_i(\text{hkl})|$, where I_i(hkl) is an individual intensity measurement and $\langle I(hkl) \rangle$ is the average intensity for this reflection

[d] R and $R_{\text{free}} = \Sigma_{\text{hkl}} ||F_\text{o}| - |F_\text{c}|| / \Sigma_{\text{hkl}} |F_\text{o}|$; for calculating R_{free}, a subset of reflections (5.0 %) was randomly chosen as a test set

[e] *DPI* Diffraction-component precision indicator (of the atom position)

[f] *RMSD* Root-mean-square deviation from ideal Engh–Huber parameters

[g] These data are relative to the native and the F-complex structures determined with the highest quality among those determined in this study
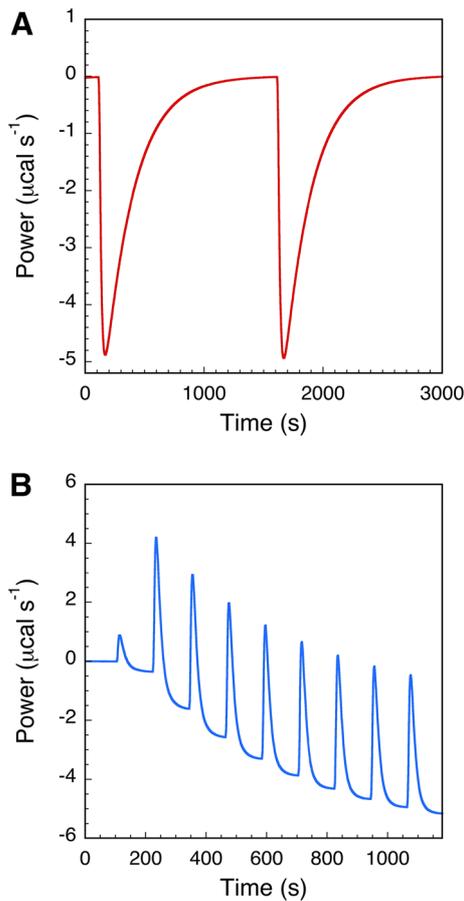
Fig. 1 Typical plots of thermal power as a function of time in M1 (**a**) and M2 (**b**) experiments. Experimental conditions: 30 nM SPU, 0.1 mM urea, 20 mM HEPES, 50 mM Na_2SO_3, 150 mM NaCl, pH 7.0

required to maintain isothermal conditions for the exothermic reaction (Fig. 1b). The rate of heat generated by the enzyme is equivalent to the decrease in thermal power after each injection, ever increasing as the substrate concentration increases. The reaction rate could thus be monitored as a function of substrate concentration and pH, as shown in Fig. 2. These data, indicating a progressive urease inhibition by increasing concentration of fluoride in the 0–800 μM range, were fit to Eq. (7) and yielded the kinetic parameters reported in Table 3. The increase of k_{cat} as a function of pH in the explored range parallels previous reports [42], while the Michaelis–Menten constant, similar to the value of 17.3 mM reported in phosphate buffer [42, 43] is invariant with pH. Tables 4 and 5 also reports the values obtained for the competitive (K_{ic}) and uncompetitive (K_{iuc}) inhibition constants: K_{iuc} is generally smaller than K_{ic}, indicating a predominance of the uncompetitive inhibition within the
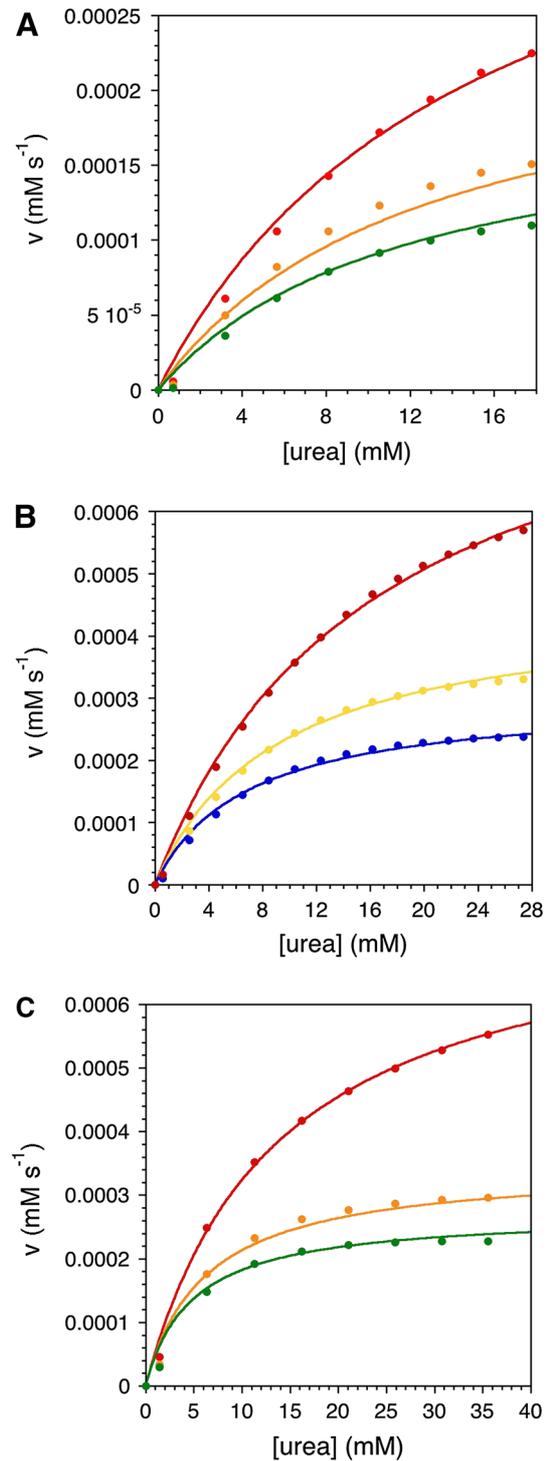
Fig. 2 Michaelis–Menten plots of *S. pasteurii* urease activity as a function of fluoride concentration (*red dots* 0 μM, *orange dots* 300 μM, *yellow dots* 400 μM, *green dots* 500 μM, *blue dots* 800 μM) at pH 6.5 (**a**), pH 7.0 (**b**) and pH 8.0 (**c**). The *lines* represent calculated non-linear regression analysis fits carried out using Eq. (7). Experimental conditions described in "Materials and methods"

Table 3 Kinetic parameters for the inhibition of SPU with fluoride

pH	k_{cat} (s^{-1})	K_M (mM)	K_{ic} (mM)	K_{iuc} (mM)
6.5	$(3.90 \pm 0.24) \times 10^3$	15.0 ± 2.0	0.79 ± 0.26	0.43 ± 0.09
7.0	$(9.01 \pm 0.20) \times 10^3$	16.5 ± 0.7	3.00 ± 0.79	0.39 ± 0.02
8.0	$(1.57 \pm 0.19) \times 10^4$	13.0 ± 1.0	–	0.28 ± 0.02

explored pH range. In addition, an increase of the uncompetitive inhibition mechanism upon pH increase is observed, while the contribution from competitive inhibition decreases and becomes negligible at pH 8.0.

Crystallographic analysis of the protein structure

The protein structure was determined independently for five crystals of the native enzyme and for five crystals of urease obtained in the presence of fluoride in the crystallization buffer. The crystals were selected using the same conditions but different hanging drops from independent crystallization trials. In the following discussion, the structures of the highest quality datasets deposited in the PDB are described unless explicitly stated.

The structure of wild-type *S. pasteurii* urease in the native form has been determined with the highest resolution so far available (1.58 Å) for ureases from bacteria or plant origin [5], resulting in more precise structural parameters, while the structure of the enzyme obtained in the presence of fluoride has been refined using data at 1.59 Å resolution, thus allowing us to analyze the effect of the inhibitor on the fine structural details of the enzyme active site. The two structures show the well-known heteropolymeric nature of this enzyme, with its threefold symmetric $(\alpha\beta\gamma)_3$ quaternary structure (Fig. 3), an α-subunit made of an $(\alpha\beta)_8$ TIM barrel domain and a β-type domain, a β-subunit characterized by numerous β strands, and a γ-subunit built as a two-layer αβ sandwich. The structure of the enzyme obtained in the presence of fluoride closely matches that of the native protein, with a global pairwise root-mean-square deviation $\left(\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \delta_i^2} \right)$, where δ is the distance between N pairs of equivalent atoms) per residue between the backbone atoms of the two structures of only 0.18, 0.06, and 0.06 Å for the α, β, and γ subunits, respectively, and an overall RMSD of 0.16 Å. The analysis of the residue-averaged backbone B-factors (Fig. 4) confirms, as previously observed [26], the presence of two characteristic helix-loop-helix regions in the α-subunit that feature a significantly larger mobility than the rest of the protein: one of these (region A, residues 375–405, Fig. 5a) is located on a surface patch that has been proposed [26] to be implicated in the formation of protein–protein complexes with the accessory

proteins that form the supercomplex necessary for the incorporation of the Ni(II) ions into the active site and consequent urease activation [5, 44], while the other (region B, residues 310–340, Fig. 5b) is the flap that modulates substrate access to the active site cavity. Consistently with their larger backbone mobility, the analysis of the global residue backbone root-mean-square deviation (RMSD) reveals significant differences in these two regions between the structures of the enzyme obtained in the absence and presence of fluoride, with the active site flap found in the open conformation in both cases, but slightly more closed in the case of latter structure (Fig. 5b). So far, only in the case of urease inhibited by diamidophosphate (DAP), an analog of the catalytic intermediate or transition state, the *S. pasteurii* active site urease flap has been found in the fully closed conformation [13]. Plots of the difference in ϕ and ψ protein backbone angles (Fig. 4) show a relatively small modification of these parameters in the active site flap, indicating that the residues in region B move all together consistently, leaving unaltered the inner structure of this motif (Fig. 5b). On the other hand, large changes are observed in region A, coherently with the observed largely modified structure of the loop between the two helices (Fig. 5a). In addition, the plot of RMSD per residue (Fig. 4) also reveals another protein segment that undergoes significant changes in the presence of fluoride: this is the loop comprising residues 365–368 in the α subunit, located in close proximity to the Ni(II) ions in the active site (Fig. 5c), and within which a large change of ϕ (113.7°) and ψ (97.9°) for αAla366 is observed. With this exception, the positions of the conserved amino acid residues that are not involved in nickel binding but are thought to be important in the catalytic mechanism (αAla170, αHis222, αGlu223, αAsp224, αGly280, αHis323, and αMet367) are largely invariant between the native and fluoride-bound urease.

Crystallographic analysis of the active site

Fluoride ions are isoelectronic with water molecules and therefore in crystallographic structures the expected heights of electron density peaks are very similar. Moreover, the radii of fluoride (1.36 Å) and water (1.40 Å) are so similar that they cannot be used alone to identify one ligand over the other. The main difference between fluoride and water (or hydroxide ions) rests in the capability of the latter, and not of fluoride, to act as hydrogen-bonding donor. Therefore, the identification of fluoride ions in protein structures must rely on either very-high-resolution crystallographic data, and/or upon other criteria, such as differences in the overall protein structural environment around the putative fluoride-binding site, in addition to other evidence derived from non-crystallographic data.

Table 4 Selected distances and angles around the Ni(II) ions in the five crystals of native SPU

	Native 4CEU	Crystal native #2	Crystal native #3	Crystal native #4	Crystal native #5	Average	Standard deviation
Ni–L distances (Å)							
Ni1–αLys220* Oθ1	1.94	1.96	1.96	1.95	1.95	1.95	0.01
Ni1–L$_B$[a]	2.08	2	2	1.98	2.01	2.01	0.03
Ni1–L1[b]	2.24	2.21	2.2	2.14	2.21	2.20	0.03
Ni1–αHis249 Nδ	2.03	2.03	2.05	2.03	2.03	2.03	0.01
Ni1–αHis275 Nε	2.02	2.03	2.01	2.04	2.03	2.03	0.01
Ni2–αLys220* Oθ2	2.08	2.02	2	2.02	2.01	2.03	0.03
Ni2–L$_B$	2.12	2.05	2.04	2.03	2.04	2.06	0.04
Ni2–L2[c]	2.07	2.14	2.07	2.06	2.14	2.10	0.04
Ni2–αHis137 Nε	2.11	2.07	2.07	2.06	2.06	2.07	0.02
Ni2–αHis139 Nε	2.08	2.06	2.06	2.05	2.07	2.06	0.01
Ni2–αAsp363 Oδ1	2.10	2.05	2.06	2.06	2.04	2.06	0.02
Ni1–Ni2	3.67	3.61	3.61	3.56	3.61	3.61	0.04
L1–L2	2.37	2.35	2.35	2.28	2.33	2.34	0.03
L–Ni–L angles (°)							
αLys220* Oθ1–Ni1–αHis249 Nδ	100.4	99.3	99.0	99.5	99.5	99.5	0.5
αLys220* Oθ1–Ni1–αHis275 Nε	107.2	107.7	107.6	106.4	106.7	107.1	0.5
αLys220* Oθ1–Ni1–L$_B$	96.6	93.7	94.1	93.3	92.9	94.1	1.5
αLys220* Oθ1–Ni1–L1	108.2	107.5	108.0	108.0	108.2	108.0	0.3
αHis249 Nδ–Ni1–αHis275 Nε	98.6	96.0	96.5	95.5	96.4	96.6	1.2
αHis275 Nε–Ni1–L$_B$	94.6	95.2	94.9	92.5	94.6	94.3	1.1
L$_B$–Ni1–L1	67.0	72.6	68.8	75.6	72.4	71.3	3.4
L1–Ni1–αHis249 Nδ	89.3	88.2	89.9	88.8	89.1	89.1	0.6
αHis249 Nδ–Ni1–L$_B$	154.2	159.5	157.7	162.4	160.3	158.8	3.1
αHis275 Nε–Ni1–L1	141.6	143.4	142.2	144.1	143.1	142.9	1.0
αLys220* Oθ2–Ni2–αHis137 Nε	90.8	91.3	89.5	90.5	91.0	90.6	0.7
αLys220* Oθ2–Ni2–αHis139 Nε	91.7	91.7	92.0	91.3	92.1	91.8	0.3
αLys220* Oθ2–Ni2–L2	92.9	94.1	95.0	96.3	94.4	94.5	1.2
αLys220* Oθ2–Ni2–L$_B$	95.6	93.9	94.3	94.6	93.4	94.4	0.8
αAsp363 Oδ1–Ni2–αHis137 Nε	82.8	82.6	82.1	82.9	82.7	82.6	0.3
αAsp363 Oδ1–Ni2–αHis139 Nε	86.4	85.6	84.7	85.2	85.0	85.4	0.7
αAsp363 Oδ1–Ni2–L2	94.5	93.1	94.8	91.9	93.1	93.5	1.2
αAsp363 Oδ1–Ni2–L$_B$	89.1	91.2	92.5	91.2	91.9	91.2	1.3
L2–Ni2–L$_B$	67.7	72.2	71.1	75.1	73.7	71.9	2.8
L$_B$–Ni2–αHis137 Nε	95.0	91.4	93.6	86.8	91.1	91.6	3.1
αHis137 Nε–Ni2–αHis139 Nε	108.5	108.1	108.2	110.1	107.9	108.6	0.9
αHis139 Nε–Ni2–L2	88.4	87.9	86.7	87.5	86.9	87.4	0.6
αLys220* Oθ2–Ni2–αAsp363 Oδ1	172.4	172.2	169.5	171.0	171.8	171.4	1.2
L$_B$–Ni2–αHis139 Nε	155.3	159.6	157.4	162.1	160.2	158.9	2.6
αHis137 Nε–Ni2–L2	162.6	163.0	164.3	161.1	164.1	163.0	1.3
Ni1–L$_B$–Ni2	122.1	125.6	127.0	124.8	126.2	125.1	1.9

Table 5 Selected distances and angles around the Ni(II) ions in the five crystals of fluoride-inhibited SPU

	Fluoride 4CEX	Crystal fluoride #2	Crystal fluoride #3	Crystal fluoride #4	Crystal fluoride #5	Average	Standard deviation
Ni–L Distances (Å)							
Ni1–αLys220* Oθ1	1.94	1.95	1.94	1.94	1.94	1.94	0.01
Ni1–L$_B^a$	2.02	2.02	2	2.02	2.04	2.02	0.01
Ni1–L1^b	2.08	2.09	2.04	2.07	2.08	2.07	0.02
Ni1–αHis249 Nδ	2.06	2.04	2.04	2.05	2.05	2.05	0.01
Ni1–αHis275 Nε	2.06	2	2.04	2.04	2.03	2.03	0.02
Ni2–αLys220* Oθ2	2.06	2.02	2.01	2.01	2.02	2.02	0.02
Ni2–L$_B$	2.01	2.04	2.08	2.04	2.03	2.04	0.03
Ni2–L2^c	2.11	2.08	2.03	2.07	2.08	2.07	0.03
Ni2–αHis137 Nε	2.11	2.07	2.05	2.04	2.07	2.07	0.03
Ni2–αHis139 Nε	2.08	2.06	2.06	2.06	2.06	2.06	0.01
Ni2–αAsp363 Oδ1	2.14	2.06	2.05	2.05	2.04	2.07	0.04
Ni1–Ni2	3.53	3.55	3.54	3.54	3.54	3.54	0.01
L1–L2	2.53	2.4	2.54	2.48	2.47	2.48	0.06
L–Ni–L angles (°)							
αLys220* Oθ1–Ni1–αHis249 Nδ	99.1	99.2	99.2	98.9	99.6	99.2	0.3
αLys220* Oθ1–Ni1–αHis275 Nε	107.1	106.2	105.5	106.3	106.3	106.3	0.6
αLys220* Oθ1–Ni1–L$_B$	92.6	91.7	93.9	93.4	92.7	92.9	0.8
αLys220* Oθ1–Ni1–L1	107.4	107.1	108.7	107.9	108.4	107.9	0.7
αHis249 Nδ–Ni1–αHis275 Nε	92.4	93.9	94.0	93.5	94.1	93.6	0.7
αHis275 Nε–Ni1–L$_B$	90.6	90.3	92.0	89.6	90.1	90.5	0.9
L$_B$–Ni1–L1	84.4	81.5	81.7	82.9	81.9	82.5	1.2
L1–Ni1–αHis249 Nδ	85.7	88.0	86.5	86.7	86.9	86.8	0.8
αHis249 Nδ–Ni1–L$_B$	166.5	166.7	164.6	166.0	165.3	165.8	0.9
αHis275 Nε–Ni1–L1	145.3	145.9	145.2	145.4	144.5	145.2	0.5
αLys220* Oθ2–Ni2–αHis137 Nε	92.1	92.3	90.5	91.0	91.7	91.5	0.8
αLys220* Oθ2–Ni2–αHis139 Nε	90.3	91.3	90.3	91.0	91.6	90.9	0.6
αLys220* Oθ2–Ni2–L2	96.1	95.1	98.6	96.8	98.2	96.9	1.5
αLys220* Oθ2–Ni2–L$_B$	95.3	95.1	96.5	95.5	96.0	95.7	0.6
αAsp363 Oδ1–Ni2–αHis137 Nε	82.5	82.7	81.2	82.5	81.8	82.1	0.6
αAsp363 Oδ1–Ni2–αHis139 Nε	84.6	84.3	85.0	84.2	85.1	84.6	0.4
αAsp363 Oδ1–Ni2–L2	90.7	91.5	91.3	91.2	89.5	90.8	0.8
αAsp363 Oδ1–Ni2–L$_B$	91.4	91.0	90.8	91.2	89.4	90.7	0.8
L2–Ni2–L$_B$	82.1	80.8	80.5	81.3	80.7	81.1	0.7
L$_B$–Ni2–αHis137 Nε	84.6	84.3	86.7	85.7	85.5	85.4	0.9
αHis137 Nε–Ni2–αHis139 Nε	108.3	110.6	108.0	107.4	108.9	108.6	1.2
αHis139 Nε–Ni2–L2	84.2	83.9	84.0	84.8	83.9	84.2	0.4
αLys220* Oθ2–Ni2–αAsp363 Oδ1	171.0	171.7	168.6	170.3	171.2	170.6	1.2
L$_B$–Ni2–αHis139 Nε	165.7	163.8	163.8	165.3	163.6	164.4	1.0
αHis137 Nε–Ni2–L2	165.0	163.6	165.0	165.5	163.7	164.6	0.8
Ni1–L$_B$–Ni2	121.9	122.2	120.0	121.3	120.8	121.1	0.9

[a] L$_B$ indicates the atom bridging Ni1 and Ni2

[b] L1 indicates the terminal atom bound to Ni1, interpreted as water for the native enzyme and fluoride for the inhibited enzyme

[c] L2 indicates the terminal atom bound to Ni2, interpreted as water for both the native and the fluoride complex
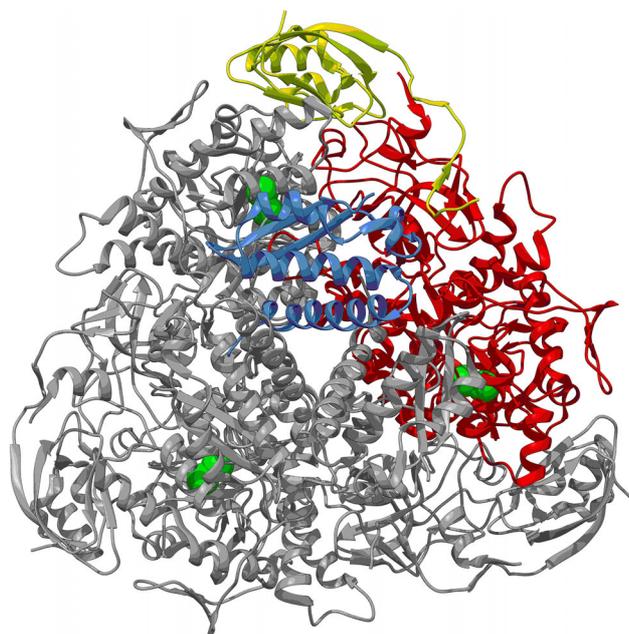
Fig. 3 Ribbon scheme of native *S. pasteurii* urease; the α, β and γ subunits of one (αβγ) trimer are shown in *red*, *yellow* and *blue*, while the positions of the Ni(II) ions in the active sites are shown as *green spheres*

In this case, clear and macroscopic differences are consistently observed between the two sets of SPU structures in the absence and presence of fluoride in the crystallization solution. First, a large change consistently involves residues 365–368 (Fig. 5c): in this loop the αAla366 residue modifies its conformation from pointing the backbone carbonyl C=O group away from the active site Ni(II) ions in the native enzyme to pointing towards them in the fluoride-inhibited complex, forming an H-bond with the terminal ligand of Ni2, which must thus carry a hydrogen atom, excluding the presence of fluoride in this position. Only in the case of the enzyme complexed with DAP [13] or phosphate [24] was the αAla366 residue observed to exist in the same conformation as found in the fluoride-inhibited urease, indicating that the presence of fluoride has analogously altered the active site structure of the native enzyme. A second macroscopic structural difference between the two sets of structures is the consistent presence of a sulfate ion in the vicinity of the nickel active site in the case of native urease, while this moiety is absent in the case of the enzyme crystallized in the presence of fluoride (Figs. 6, 7). This is a strong argument in favor of the binding of fluoride ions at the active site because the increased negative charge brought by the latter would repel the negative charge of the sulfate ion and alter the hydrogen-bonding network that stabilizes the sulfate in that position.

In addition to these two macroscopic structural differences, finer modifications of distances and angles are observed between the two highest quality structures (Tables 4, 5), whose significance is discussed here below. In the case of native SPU, the electron density in the active site is well defined, as shown in Fig. 6a; the corresponding structural model is presented in Fig. 7a. The Ni(II) ions are ordered (B-factors of 9.67 and 8.65 Å^2 for Ni1 and Ni2, respectively) and separated by 3.67 Å. They are bridged by a ligand, interpreted as a hydroxide ion, W_B [13], with distances to Ni1 and Ni2 of 2.08 and 2.12 Å, respectively, and by the carboxylate group of the carbamylated αLys220* (Ni1–Oθ1 = 1.94 Å; Ni2–O02 = 2.08 Å). The B-factors of the bridging (13.75 Å^2) and terminal (13.99 and 11.98 Å^2) solvent Ni(II) ligands are consistent with full occupancy. Ni1 is further coordinated by αHis249 Nδ (at 2.03 Å) and by αHis275 Nε (at 2.02 Å), while Ni2 is bound to αHis137 Nε (at 2.11 Å), αHis139 Nε (at 2.08 Å), and αAsp363 Oδ1 (at 2.10 Å) (Fig. 7a). The two Ni(II) ions are additionally bound to a terminal ligand, interpreted as a water molecule [13], (Ni1–W1 = 2.24 Å; Ni2–W2 = 2.07 Å), making Ni1 penta-coordinated in a tetragonal pyramidal geometry and Ni2 hexa-coordinated and pseudo-octahedral. A fourth distal solvent molecule (W3) completes a tetrahedral cluster of electron density peaks in the urease active site, with distances consistent with the formation of hydrogen bonds (W3–W1 = 2.33 ; W3–W2 = 2.37 Å; W3–W_B = 2.47 Å) and a B-factor that indicates full occupancy (15.72 Å^2).

In the case of SPU complexed with fluoride (Figs. 6b, 7b for the highest quality structure of this set), a series of structural modifications are observed (Tables 4, 5). To statistically analyze these differences it is necessary to estimate the standard deviations of the atomic positional parameters, or estimated standard uncertainties (ESU). The currently most accepted method to perform this analysis has been proposed by Cruickshank [45]. In this method, the uncertainties [diffraction-component precision indicator (DPI)] are estimated based on atom type and number, average B-value, as well as data completeness, crystallographic R-factor and resolution. These parameters are reported in Tables 1 and 2, together with additional average-based positional error estimates, which are less restrictive as compared to the analysis carried out below, based on DPI. First of all, the Ni–Ni distance decreases from 3.67 to 3.53 Å, a difference significant with a confidence level of 90.7 % (using $\sigma_{12} = 2 \cdot DPI$, and $Z = \Delta_{distance}/\sigma_{12}$). A decrease is also observed for the distance between the penta-coordinated Ni1 and its terminal ligand (from 2.24 to 2.08 Å), statistically significant at the 94.5 % confidence level. This modification is
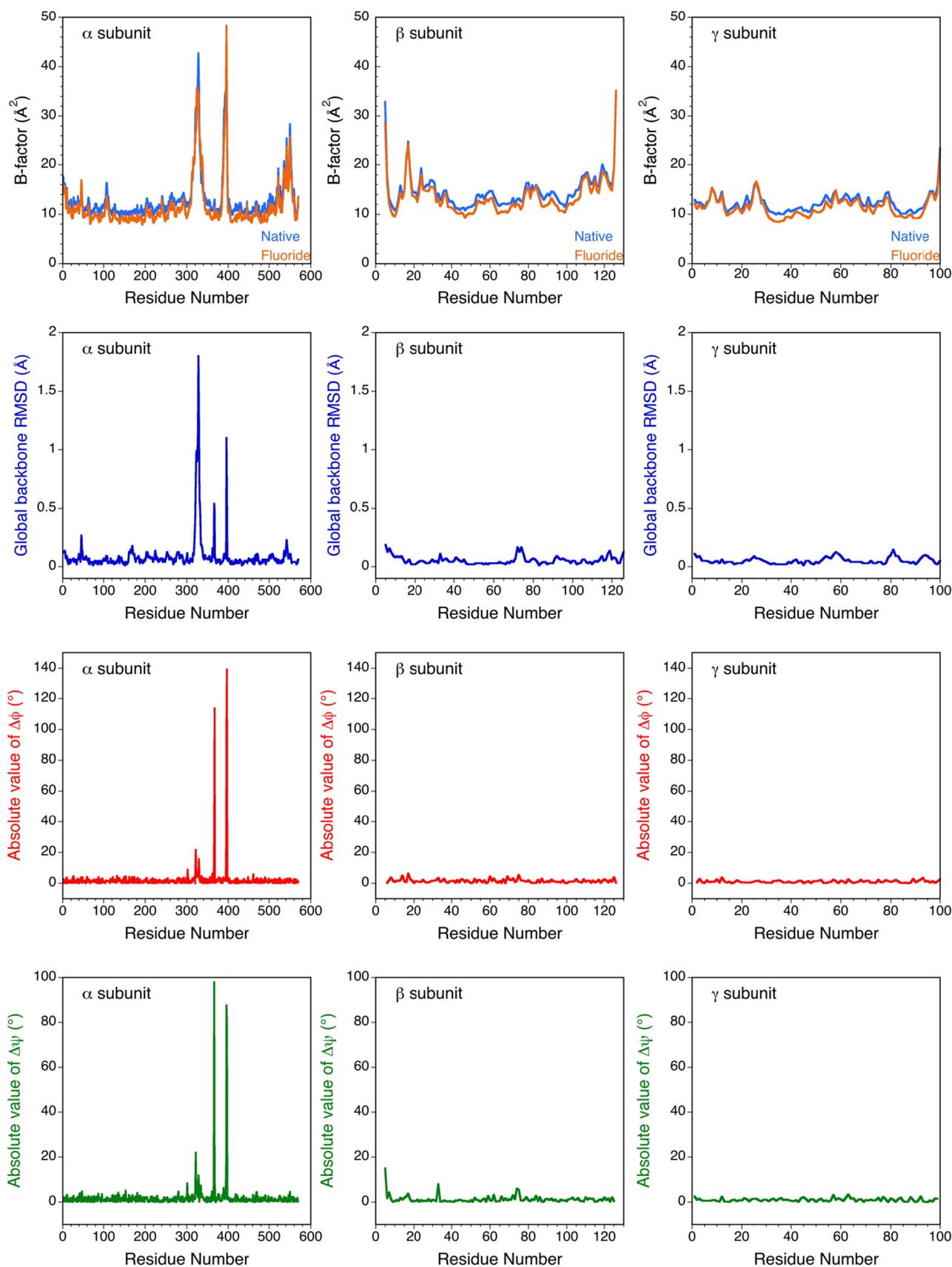
◄ **Fig. 4** Comparison of the structural parameters between the α, β and γ subunits of native (PDB code 4CEU) and fluoride-bound (PDB code 4CEX) *S. pasteurii* urease: global pairwise RMSD per residue of the backbones (*top panels*), residue-averaged backbone B-factors (*middle panels*), and absolute values of the $\Delta\phi$ and $\Delta\psi$ protein backbone angles (*bottom panels*)

accompanied by the decrease from 2.84 to 2.72 Å (statistic significance of 85.0 %) of the distance between the protonated αHis222 Nε and the atom terminally bound to Ni1, an interaction that involves a hydrogen bond. The distance between the bridging ligand and the nickel ions decreases from 2.08 and 2.12 Å in native SPU to 2.02 and 2.01 Å for Ni1 and Ni2, respectively. These differences correspond to a statistic significance of 52.9 and 81.3 % confidence level. The distance between the hexa-coordinated Ni2 and its terminal ligand increases from 2.07 to 2.11 Å (36.9 % significance level).

Several additional modifications are observed in the structure of SPU obtained in the presence of fluoride. In particular, the distance between the distal water molecule W3 (slightly disordered or not fully occupied, as indicated by its B-factor = 26.93 $Å^2$) and the three other centers that compose the tetrahedron of electron density maxima, interpreted as solvent molecules in the structure of native urease, increases significantly (3.20 vs. 2.33 Å, 3.54 vs. 2.37 Å, and 3.05 vs. 2.47 Å). Moreover, while in the case of the native enzyme W3 forms a hydrogen bond with αGly280 at 3.52 Å, in the structure of urease in the presence of fluoride this hydrogen bond measures only 2.69 Å (Fig. 7).

The similarity of the radii of fluoride and water or hydroxide cannot be used to identify the metal ligand based only on bond distances. However, the scattering factor of fluoride is generally larger at the resolution corresponding to the limit of our data (International Tables for X-ray Crystallography). Therefore, we carried out an alternative refinement of the crystallographic data collected on the crystal of SPU obtained in the presence of fluoride using a model that contained oxygen atoms terminally bound to Ni1 and in the nickel-bridging position. This refinement provided a B-factor value for the bridging atom significantly smaller (5.88 $Å^2$) than that determined using fluoride in this position (8.28 $Å^2$), and even smaller than that of the nickel atoms (9.67 and 8.75 $Å^2$ for Ni1 and Ni2, respectively), an unrealistic result that strongly supports the presence of fluoride in the bridging position. This refinement further yielded a B-factor value of 10.91 $Å^2$ for the oxygen atom bound to Ni1, suggesting the presence of fluoride (B-factor of 13.60 $Å^2$), but not ruling out, if based only on this calculation, partial occupancy of fluoride and water in this position. The differences between the two highest

quality datasets, described above, are further supported by the analysis of the distances determined for two independent sets, each made of five crystals of urease obtained, respectively, in the absence and presence of fluoride in the crystallization solution (Tables 4, 5). The average Ni1–Ni2 distance in the native enzyme is 3.61 ± 0.04 Å, while in the presence of fluoride a distance of 3.54 ± 0.01 Å is observed; the Ni1–L1 average distances are 2.20 ± 0.03 vs. 2.07 ± 0.02 Å, the Ni2–L2 average distances are 2.10 ± 0.04 vs. 2.07 ± 0.03 Å, and the L1–L2 average distances are 2.34 ± 0.03 vs. 2.48 ± 0.06 Å.

In addition to bond distances, some modifications between the two structures also involve bond angles, with the largest changes involving the terminal and bridging ligands of the Ni(II) ions (see Tables 4, 5). In particular, the geometry around the penta-coordinated Ni1 in the two structures can be compared using the structural τ parameter based on Scheme 3, shown below:

Scheme 3 Geometry around the penta-coordinated Ni1
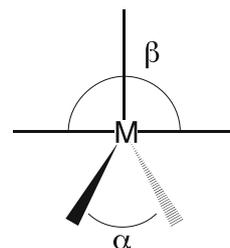
where $\tau = \frac{(\beta-\alpha)}{60}$, and α and β (β ≥ α) are the two largest angles around the metal ion. τ = 1 for a perfectly trigonal bipyramidal (TB) geometry, whereas τ = 0 for a square pyramidal geometry (SPY) [46]. The Ni1 atom in the native urease has τ = 0.21, while that observed for the fluoride-bound enzyme has τ = 0.35, indicating a distortion from the mainly SPY geometry to a more TB coordination. Tables 4 and 5 further report a comparison of bond angles around the Ni(II) ions among all structures determined in this study, from which significant differences are found mainly in the case of angles involving the L_B, L1 and L2 positions. The picture emerging from this analysis is somehow consistent with the interpretation of the electron density map obtained at 2.2 Å resolution for the complex of JBU with fluoride (PDB code 4GOA) as containing a nickel-bridging fluoride, while an additional F^- ion was assigned in the position occupied by W3 in the native enzyme. In that case, it is not clear, in the absence of a published report, how the model could be accurately refined at this resolution.

In conclusion, in the presence of fluoride the structure of the nickel environment in the enzyme active site has been
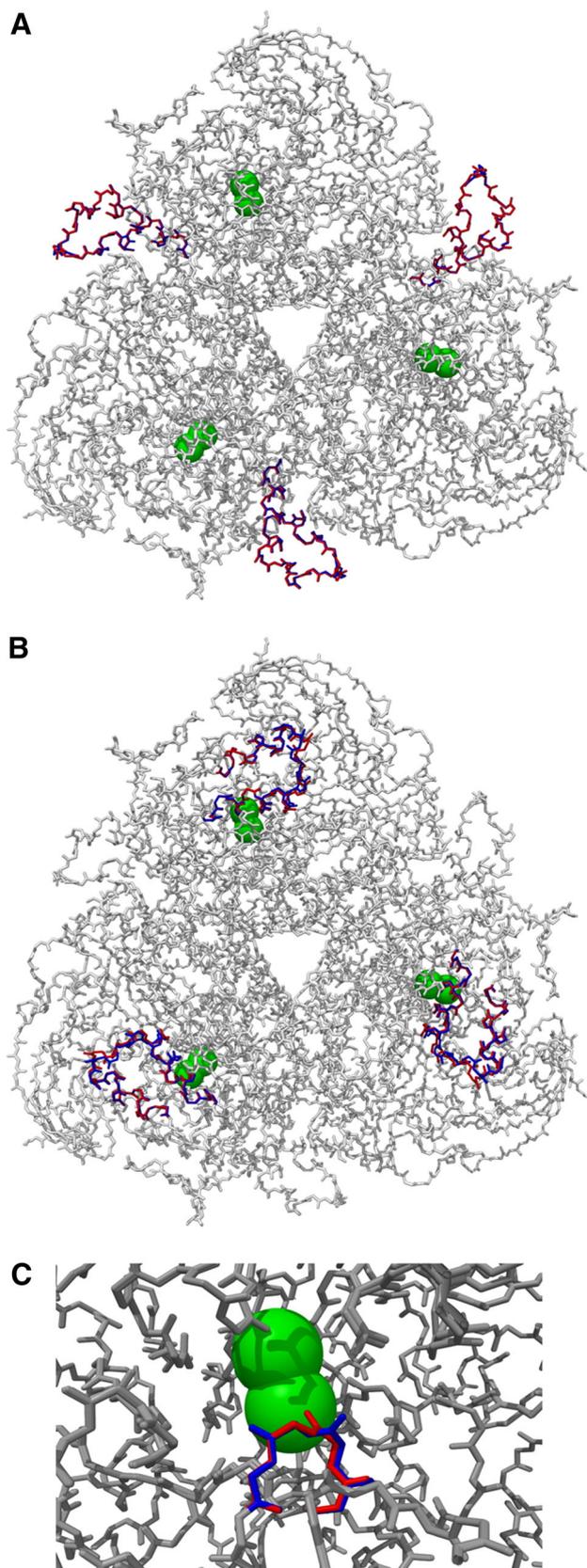
◄**Fig. 5** Structure of the backbone mobile regions of *S. pasteurii* urease in the native (*blue*) and in the fluoride-inhibited (*red*) form: **a** residues 375–405; **b** residues 310–340; **c** residues 365–368

significantly altered. The largest modifications involve the rearrangement of the protein backbone in the proximity of the active site and the removal of molecules of water and sulfate from the enzymatic cavity. In addition, statistically significant distances and angles that feature non-protein-based ligands are observed, highlighting on one hand the substantial rigidity of the enzyme on nickel-coordination sites defined by the protein framework, and on the other hand the flexibility of active cavity positions occupied by solvent molecules in the native form.

Discussion

The first report on urease inhibition by fluoride dates back to 1928, when Mystokowski revealed that NaF is a powerful inhibitor of ordinary urease [47], followed by a study by Pearson and Smith in 1943, in which the effect of fluoride on the activity of bovine rumen urease was described [48]. Following their first evidence of the essentiality of nickel for urease activity in 1975 [49] and in the absence of any structural information yet derived from crystallography, Dixon, Blakeley and Zerner reported in 1980 pioneering and insightful experiments that indicated that fluoride is a inhibitor of jack bean urease (JBU) by virtue of binding to the active site Ni(II) ions [27]. The support for this proposal was the similarity of the dissociation constant of a Ni(II)–fluoride complex (1.23 ± 0.10 mM) determined by monitoring the metal-based electronic transitions by visible spectroscopy, and the kinetically determined fluoride competitive inhibition constant, K_{ic}, at pH 7. The latter corresponded to 1.01 ± 0.10 mM for the instantaneously formed fluoride–urease complex (pre-steady state), and to 0.83 ± 0.27 mM for an analogous complex formed overtime (steady state) with a rate of formation of 0.53–2.3 min^{-1} depending on fluoride and urea concentrations [27]. The authors further proposed that this time-dependent process involved the formation of a ternary complex of JBU–fluoride–urea (or JBU–fluoride–carbamate), for which the apparent dissociation constant of the fluoride ion is ca. 0.23 mM [27]. The authors did not explicitly use the term "uncompetitive inhibition", even though they reported double-reciprocal plots at steady state that showed parallel lines, consistent with this type of inhibition. With the knowledge of the presence of two Ni(II) ions in the active site, but in the
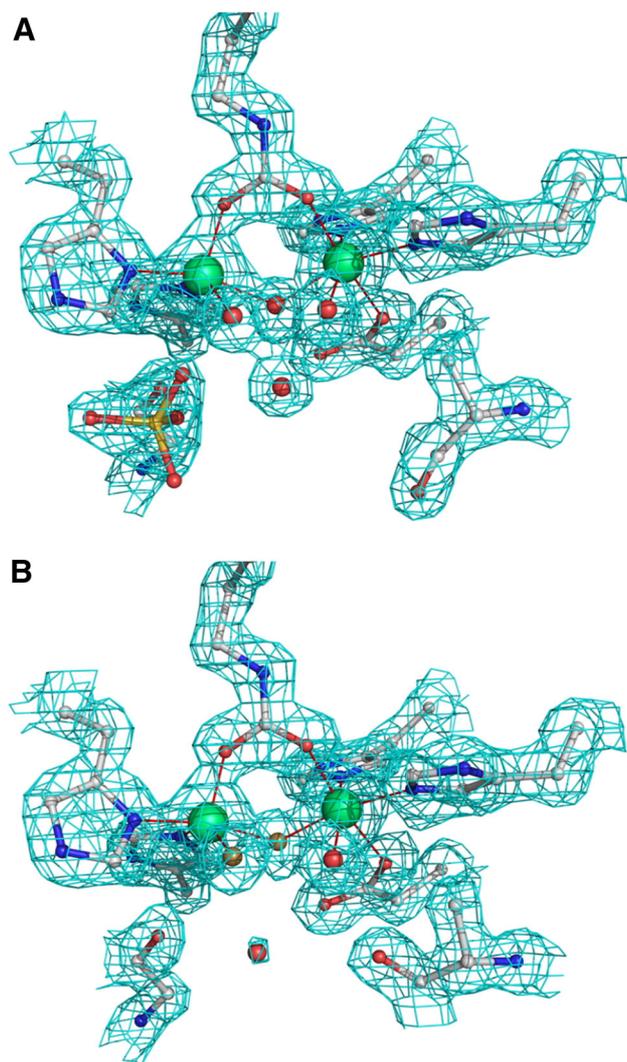
Fig. 6 Atomic model of the active site of native (panel **a**, PDB code 4CEU) and fluoride-bound (panel **b**, PDB code 4CEX) *S. pasteurii* urease. The nickel-coordination environment is shown superimposed on the final *2Fo–Fc* electron density map contoured at 1.7 σ. The carbon, nitrogen, oxygen, fluorine, sulfur and nickel atoms are *grey*, *blue*, *red*, *gold*, *yellow* and *green*, respectively

Fig. 7 Crystallographic structural model for the active site obtained for native (panel **a**, PDB code 4CEU) and fluoride-bound (panel **b**, PDB code 4CEX) *S. pasteurii* urease. Figures created using Crystal-Maker. The carbon, nitrogen, oxygen, fluorine and nickel atoms are *grey*, *blue*, *red*, *gold* and *green*, respectively. Hydrogen bonds are shown as thin *blue lines*. *Spheres* are drawn using the relative atomic radii values in CrystalMaker

absence of any structural information on the enzyme, the authors acknowledged that the kinetic methods that were used could not indicate whether fluoride and urea (or carbamate) bind to the same or to different nickel ions [27]. In 1997, Saboury and Moosavi-Movahedi described the application of isothermal titration calorimetry for the determination of the kinetic enzymatic parameters for JBU [29]. In particular, the effect of fluoride on the enzyme kinetics was reported as being competitive, with $K_{ic} = 0.94$ mM. In 2000, Todd and Hausinger investigated the inhibition of urease from *Klebsiella aerogenes* (KAU) by fluoride, and concluded that KAU is slowly inhibited by this anion (and not its protonated form HF) in both the presence and absence of urea, consistently with a non-

classical uncompetitive slow-binding inhibition mechanism [28]. In particular, the time constant for the slow inhibition process, complete in about 5 min, ranged between 0.26 and 0.98 min^{-1}, depending on fluoride concentration (in the 2.5–10 mM range), similarly to what reported by Dixon et al. [27] for JBU. The value of K_{iuc} at pH 7, estimated using a plot of the pH dependence of this parameter reported in the article, is ca. 0.16 mM. The authors also proposed that fluoride binds to a form of the enzyme generated during the catalytic process, because greater inhibition is observed in the presence of substrate [28]. The

onset of uncompetitive inhibition is usually observed for multi-substrate enzymes, and on this basis the authors suggested that fluoride binds to a site different from that used for urea binding and occupied instead by the co-substrate, that is, the catalytic water molecule. The authors interpreted these results using the structure of urease by then available from different bacterial sources [12, 13, 21]. In particular, they speculated that fluoride binds to the urease metallocenter with displacement of the nickel-bridging hydroxide (W_B) and excluded substitution of W1 or W2, replaced by urea during catalysis, because no competitive inhibition was detected [28]. Finally, in 2001 Krajewska et al. [30] reported kinetic data on JBU interpreted with a model in which fluoride binds to urease in a competitive slow-binding mechanism, consisting of an initial rapid formation of a urease–fluoride complex having a dissociation constant $K_{ic} = 1.0$ mM, that slowly evolves to a more stable complex with a dissociation constant of 0.02 mM.

In the present study, enzymatic measurements of fluoride inhibition of urease were carried out on SPU using a calorimetric assay, and the results were corroborated thorough the analysis of high-resolution crystallographic data of native and fluoride-inhibited urease. The experimental conditions of the enzymatic assay covered the 6.5–8.0 pH range. SPU was incubated in the presence of different fluoride concentrations, with no difference observed on the kinetic parameters using either 15 or 60 min of incubation time. This ensured that the data obtained were related to the steady-state fluoride-bound enzyme, and not to any time-dependent form of the enzyme. This choice was dictated by the nature of the assay, which requires a much longer time than the few minutes necessary for the onset of the steady-state condition for JBU and KAU [27, 28].

The calorimetric data show that fluoride inhibits the activity of SPU at pH 7.0 with mixed competitive and uncompetitive mechanisms. In particular, the uncompetitive inhibition constant is ca. sevenfold lower than the competitive inhibition constant ($K_{iuc} = 0.39 \pm 0.02$ vs. $K_{ic} = 3.00 \pm 0.79$ mM), indicating the predominance of the uncompetitive effect. These values are similar to those reported by Dixon et al. ($K_{iuc} = 0.23$ mM, $K_{ic} = 0.83 \pm 0.80$ mM), by Saboury et al. ($K_{ic} = 0.94$ mM) and by Krajewska et al. ($K_{ic} = 1.0$ mM) for JBU, and by Todd et al. (K_{iuc} ca. 0.16 mM) for KAU. Furthermore, the pH dependence of the two inhibition constants indicates that while the effect of uncompetitive mechanism increases with pH, the contribution of the competitive inhibition considerably decreases to negligible levels upon increasing the pH.

These kinetic data require the determination of the fluoride-binding positions in the enzyme active site to firmly rationalize the mechanism of inhibition. The identification of fluoride ions vs. exchangeable solvent molecules bound to metal centers in metalloenzymes using protein crystallography demands very precise structural data, because structural changes are generally fractions of Å and could remain unnoticed unless the resolution is exceptionally high. X-ray structures of proteins at true atomic resolution (<1.2 Å) are still scarce, currently representing ca. 2 % of the overall Protein Data Bank. The structures of fluoride complexes of a series of metalloenzymes containing a di-metallic core have been reported for Mn(II)-containing inorganic pyrophosphatase at 1.20 Å [50] and arginase at 2.40 Å [51], for Mg(II)-containing enolase at 1.36 Å [52] and diphosphoinositol polyphosphate phosphohydrolase-1 at 1.65 Å [53], for Zn(II)/Fe(III)-containing purple acid phosphatase at 2.20 Å [54] and for Ni(II)-dependent JBU at 2.20 Å (PDB code 4GOA, unpublished).

The crystallographic analysis to a resolution better than 1.72 Å of the ten replicated structures of SPU obtained here in the absence and presence of fluoride provided refined atomic models with very small positional errors (see Table 1) allowing us to determine the statistic significance of distance differences between the two sets of structures. In general, the electron density maps clearly show that fluoride alters the active site environment surrounding the Ni(II) ions by (1) inducing a conformational change of the protein backbone in correspondence of αAla366, promoting the formation of a H-bond between its carbonyl O atom and the terminal ligand of Ni2, (2) disrupting the hydrogen bond network that holds the distal solvent molecule W3 close to W1, W2 and W_B in the native form of the enzyme, (3) preventing sulfate binding in the active site cavity, (4) decreasing the Ni1–Ni2 distance, (5) decreasing the distance between the nickel atoms and the bridging ligand, (6) decreasing the distance between Ni1 and its terminal ligand, and (7) increasing the distance between Ni2 and its terminal ligand. These modifications of the coordination environment of the metal ions in the urease active site can be interpreted as the result of the presence of a fluoride ion in the bridging position (F_B) in place of the hydroxide ion found in the native enzyme, and the substitution of the neutral water molecule W1 bound to Ni1 by a second fluoride ion (F1) (Fig. 5a, b). The increase of the uncompetitive inhibition mechanism upon pH increase was already reported and discussed in the case of KAU [28]. On the other hand, the decrease of the competitive inhibition mechanism upon raising the pH has not been observed so far in the case of urease, and could be explained by a progressive deprotonation of W1, which would lead to a more difficult substitution by fluoride. The observation in the crystal structure of a fluoride ion terminally bound to Ni1 could be due to the pH of the

crystallization solution (6.5): indeed, in these conditions, a measurable value for the competitive inhibition constant can be calculated from the fit of the kinetic data.

The presence of fluoride in the nickel-bridging position, at a metal-binding site in which the electrostatic positive charge is expected to be greatest within the active site [55], would explain the decreased Ni–Ni distance, the decrease of the Ni1–L1 distance, and the alterations of the bond angles around Ni(II) in the presence of fluoride. Similar modifications were indeed observed in the case of native and fluoride-inhibited Mg(II)-enolase determined at 1.36 Å resolution [52]. The apparent substitution of W1, and not W2, by fluoride could be rationalized by the larger positive charge expected on Ni1, which is less coordinatively saturated as compared to Ni2, thus inducing a selective reactivity of the two Ni(II) ions. Fluoride binding to Ni1 would also be stabilized by the interaction with the protonated His222 NHε2, which is indeed located at 2.84 Å in native SPU and at 2.72 Å in the case of the fluoride–urease complex. The presence of fluoride in place of W1 could explain the larger distance between Ni2 and its terminal ligand in the fluoride-inhibited enzyme, interpreted as a water molecule W2 as due to a greater electron density on the di-nickel metal cluster induced by an increase in the number of negatively charged ligands, thus decreasing the electrophilic character of the two Ni(II) ions. In turn, the larger distance between Ni2 and W2 in the fluoride-inhibited urease could explain, together with a decrease of the occupancy of the W3 solvent molecule, the change in the backbone conformation that brings the carbonyl group of αAla366 close to W2, located at a distance that, in the case of the fluoride-bound active site and not in the native enzyme, allows the formation of a hydrogen bond. The latter conformational change could also be allowed by the dislodgment of the W3 water molecule, which would increase the steric volume available for αAla366 O. If the fluoride anion, and not water, was bound to Ni2, this movement would indeed bring two atoms with partial negative charges closer to each other. This observation could also be used to further support the mechanism step that entails the chelating mode of urea binding to the di-nickel active site of urease: the effect of substituting the neutral W1 on Ni1 by an anionic (fluoride) or neutral but polarized ligand (urea) would have the effect of weakening and lengthening the Ni2–W2 bond, favoring the interaction of urea nitrogen with Ni2 upon displacement of W2, in a structure stabilized by H-bonding interactions of urea – NH2 group with the carbonyl group of αAla366. The resulting substitution of two of the four solvent molecules in the tetrahedral cluster present in the active site of native urease with two fluoride ions apparently disrupts the hydrogen-bonding network among the four solvent molecules observed in the case of the native enzyme [13], thus explaining the large change observed in the position and occupancy/mobility of the distal water molecule. Furthermore, the lack of the sulfate ion near the nickel active site in the case of SPU crystallized in the presence of fluoride could then be rationalized by an increased negative charge in the vicinity of the active site upon substitution of a neutral water and a hydroxide ion, as found in the native enzyme, by two fluoride anions.

This picture for SPU is consistent with the prediction by Dixon et al. [49] on the direct interaction of fluoride with the Ni(II) ions in the active site of JBU: fluoride bound to Ni1 competes with the substrate urea for this electrophilic metal ion, in agreement with the currently most accepted urease catalytic mechanism [1, 5, 7, 9, 10, 13]. Moreover, the suggestion by Todd and Hausinger [28] for the substitution of the bridging hydroxide by fluoride, observed in the crystal structure of inhibited SPU, is also consistent with an uncompetitive mechanism in which fluoride competes with the second substrate of the enzymatic catalysis, namely the nucleophilic hydroxide, leaving Ni1 free to bind urea. This interpretation can be summarized by stating

Scheme 4 Native and fluoride inhibited *S. pasteurii* urease active sites. The competitive inhibition of fluoride is ascribed to the L1 position, while the uncompetitive inhibition involves the LB position
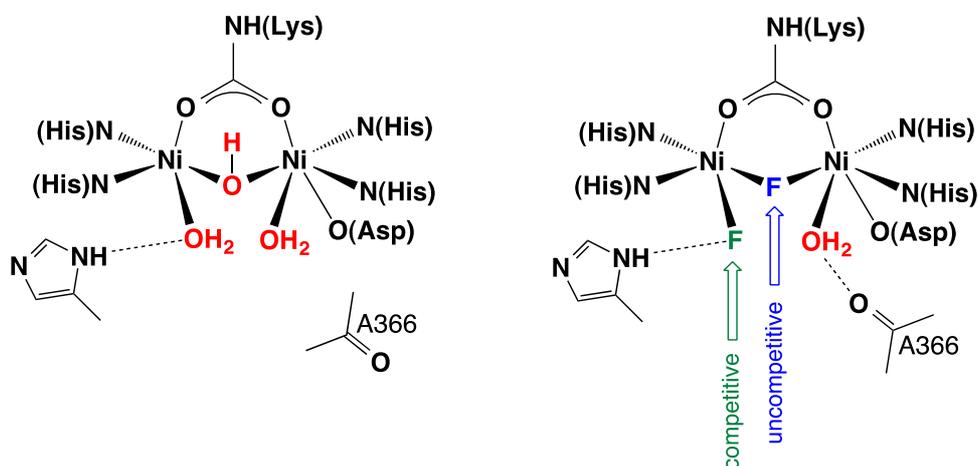
that the competitive inhibition of fluoride is ascribed to the L1 position, while the uncompetitive inhibition involves the L_B position (Scheme 4):

This is consistent with the proposed weakening of the bonds between the bridging hydroxide and the two Ni(II) ions upon urea binding to Ni1, as also implied in the proposed catalytic mechanism [1, 5, 7, 9, 10, 13]. The present study thus structurally clarifies the mechanism of urease inhibition by fluoride, and, at the same time, provides hints to understand and rationalize the role of Ni(II) in the enzymatic mechanism.

The conclusions from our study on the urease inhibition by fluoride allow us to discuss in more general terms the interaction of this anion with two other classes of binuclear metallohydrolases, namely the Mn(II)-dependent enzyme arginase [56] and the Fe(III)–M(II) [M = Fe, Zn] dependent purple acid phosphatase (PAP) [57, 58]. In the case of rat liver arginase, fluoride acts as an uncompetitive inhibitor [59], while for bovine liver arginase a non-competitive (mixed) mechanism has been reported [60]; in both cases, the inhibition decreases by increasing pH. In the case of rat liver arginase, a single fluoride appears to bind to the $[Mn(II)]_2$ cluster, while at higher pH the binding of two F^- ions was observed [51]. The 2.4 Å resolution X-ray structure of rat liver arginase was interpreted as containing two fluoride ions, one bridging the two Mn(II) ions and one terminally bound to the least coordinatively saturated Mn(II) ion [51]. In the case of pig PAP, which contains the Fe(III)–Fe(II) dinuclear cluster in the active site, fluoride inhibition is uncompetitive at pH 4.9 and lower but non-competitive (mixed) at higher pH values [61]; this pH-dependent behavior is somehow opposite to what observed for the analogous human PAP, for which the uncompetitive inhibition is dominant at pH 4.9 and above, but becomes non-competitive (mixed) at lower pH; the binding of two fluoride ions was also reported [62]. On the contrary, in the case of the Fe(III)–Zn(II) containing plant PAP the inhibition mechanism is competitive and independent of pH [61]. In all these cases, a similar order of magnitude for the inhibition constants, in the mM range, was reported, similar to the case of urease. The 2.2 Å resolution X-ray structure of red kidney bean PAP in complex with fluoride was interpreted as containing one F^- bound in the position bridging the Fe(III) and the Zn(II) ions [54]. A general trend is thus observed, with the bridging position of the dinuclear center of metallohydrolases being reactive towards fluoride substitution, while the different pH dependences might be the result of the influence of the different amino acid residues in the vicinity of the active site.

References

1. Maroney MJ, Ciurli S (2014) Chem Rev 114:4206–4228
2. Mobley HLT, Hausinger RP (1989) Microbiol Rev 53:85–108
3. Karplus PA, Pearson MA, Hausinger RP (1997) Acc Chem Res 30:330–337
4. Krajewska B (2009) J Mol Cat B Enzymatic 59:9–21
5. Zambelli B, Musiani F, Benini S, Ciurli S (2011) Acc Chem Res 44:520–530
6. Callahan BP, Yuan Y, Wolfenden R (2005) J Am Chem Soc 127:10828–10829
7. Ciurli S, Benini S, Rypniewski WR, Wilson KS, Miletti S, Mangani S (1999) Coord Chem Rev 190–192:331–355
8. Burne RA, Chen YY (2000) Microbes Infect 2:533–542
9. Ciurli S (2007) Nickel and its surprising impact in nature. Wiley, Chichester, pp 241–278
10. Benini S, Musiani F, Ciurli S (2013) In: Kretsinger RH, Uversky VN, Permyakov EA (eds) Encyclopedia of metalloproteins. Springer, New York, pp 2287–2292
11. Farrugia MA, Macomber L, Hausinger RP (2013) J Biol Chem 288:13178–13185
12. Jabri E, Carr MB, Hausinger RP, Karplus PA (1995) Science 268:998–1004
13. Benini S, Rypniewski WR, Wilson KS, Miletti S, Ciurli S, Mangani S (1999) Structure 7:205–216
14. Ha N-C, Oh S-T, Sung JY, Cha KA, Lee MH, Oh B-H (2001) Nat Struct Biol 8:505–509
15. Balasubramanian A, Ponnuraj K (2010) J Mol Biol 400:274–283
16. Balasubramanian A, Durairajpandian V, Elumalai S, Mathivanan N, Munirajan AK, Ponnuraj K (2013) Int J Biol Macromol 58:301–309
17. Kosikowska P, Berlicki Ł (2011) Expert Opin Ther Pat 21:945–957
18. Vassiliou S, Grabowiecka A, Kosikowska P, Yiotakis A, Kafarski P, Berlicki Ł (2008) J Med Chem 51:5736–5744
19. Vassiliou S, Kosikowska P, Grabowiecka A, Yiotakis A, Kafarski P, Berlicki Ł (2010) J Med Chem 53:5597–5606
20. Berlicki Ł, Bochno M, Grabowiecka A, Białas A, Kosikowska P, Kafarski P (2012) Amino Acids 42:1937–1945
21. Benini S, Rypniewski WR, Wilson KS, Ciurli S, Mangani S (1998) J Biol Inorg Chem 3:268–273
22. Pearson MA, Overbye Michel L, Hausinger RP, Karplus PA (1997) Biochemistry 36:8164–8172
23. Benini S, Rypniewski WR, Wilson KS, Miletti S, Ciurli S, Mangani S (2000) J Biol Inorg Chem 5:110–118
24. Benini S, Rypniewski WR, Wilson KS, Ciurli S, Mangani S (2001) J Biol Inorg Chem 6:778–790
25. Benini S, Rypniewski WR, Wilson KS, Mangani S, Ciurli S (2004) J Am Chem Soc 126:3714–3715
26. Benini S, Kosikowska P, Cianci M, Mazzei L, Vara AG, Berlicki Ł, Ciurli S (2013) J Biol Inorg Chem 18:391–399
27. Dixon NE, Blakeley RL, Zerner B (1980) Can J Biochem 58:481–488
28. Todd MJ, Hausinger RP (2000) Biochemistry 39:5389–5396
29. Saboury AA, Moosavi-Movahedi AA (1997) J Enz Inhib 12:273–279
30. Krajewska B, Zaborska W, Leszko M (2001) J Mol Catal B 14:101–109

31. Todd MJ, Gomez J (2001) Anal Biochem 296:179–187
32. Mazzei L, Ciurli S, Zambelli B (2014) J Vis Exp (86):e51487. doi:10.3791/51487
33. Segel IH (1993) Enzyme kinetics: behavior and analysis of rapid equilibrium and steady-state enzyme systems. Wiley-Interscience, New York
34. Kabsch W (2010) Acta Crystallogr D Biol Crystallogr 66:125–132
35. N. Collaborative Computational Project (1994) Acta Cryst D50:760–763
36. Murshudov GN, Vagin AA, Dodson EJ (1997) Acta Crystallogr D Biol Crystallogr 53:240–255
37. Murshudov GN, Vagin AA, Lebedev A, Wilson KS, Dodson EJ (1999) Acta Crystallogr D Biol Crystallogr 55:247–255
38. Emsley P, Cowtan K (2004) Acta Crystallogr D Biol Crystallogr 60:2126–2132
39. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Acta Crystallogr D Biol Crystallogr 66:486–501
40. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) J Comput Chem 25:1605–1612
41. Pedroso MM, Ely F, Lonhienne T, Gahan LR, Ollis DL, Guddat LW, Schenk GG (2014) J Biol Inorg Chem 19:389–398
42. Ciurli S, Marzadori C, Benini S, Deiana S, Gessa C (1996) Soil Biol Biochem 28:811–817
43. Benini S, Gessa C, Ciurli S (1996) Soil Biol Biochem 28:819–821
44. Quiroz-Valenzuela S, Sukuru SC, Hausinger RP, Kuhn LA, Heller WT (2008) Arch Biochem Biophys 480:51–57
45. Cruickshank DWJ (1999) Acta Cryst D55:583–601
46. Addison AW, Rao TN, Reedijk J, van Rijn J, Verschoor GC (1984) J Chem Soc Dalton Trans (7):1349–1356
47. Mystkowski EM (1928) Acta Biol Exp. Varsovie 2:212–224
48. Pearson RM, Smith J (1943) Biochem J 37:153–164
49. Dixon NE, Gazzola C, Blakeley R, Zerner B (1975) J Am Chem Soc 97:4131–4132
50. Samygina VR, Moiseev VM, Rodina EV, Vorobyeva NN, Popov AN, Kurilova SA, Nazarova TI, Avaeva SM, Bartunik HD (2007) J Mol Biol 366:1305–1317
51. Cama E, Pethe S, Boucher JL, Han S, Emig FA, Ash DE, Viola RE, Mansuy D, Christianson DW (2004) Biochemistry 43:8987–8999
52. Qin J, Chai G, Brewer JM, Lovelace LL, Lebioda L (2006) Biochemistry 45:793–800
53. Thorsell A-G, Persson C, Grslund S, Hammarstrm M, Busam RD, Hallberg BM (2009) Proteins Struct Funct Bioinform 77:242–246
54. Schenk G, Elliott TW, Leung E, Carrington LE, Mitic N, Gahan LR, Guddat LW (2008) BMC Struct Biol 8:6
55. Musiani F, Arnofi E, Casadio R, Ciurli S (2001) J Biol Inorg Chem 6:300–314
56. Christianson DW (2005) Acc Chem Res 38:191–201
57. Mitić NN, Smith SJS, Neves AA, Guddat LW, Gahan LRL, Schenk GG (2006) Chem Rev 106:3338–3363
58. Schenk GG, Mitić NN, Gahan LR, Ollis DL, McGeary RP, Guddat LW (2012) Acc Chem Res 45:1593–1603
59. Pethe SS, Boucher JLJ, Mansuy DD (2002) J Inorg Biochem 88:397–402
60. Xie X-Y, Wang C-X, Wang Z-Y (2004) J Therm Anal Calorim 77:1005–1012
61. Elliot TW, Mitić NN, Gahan LR, Guddat LW, Schenk GG (2006) J Braz Chem Soc 17:1558–1565
62. Mitić NN, Valizadeh M, Leung EW, de Jersey J, Hamilton S, Hume DA, Cassady AI, Schenk GG (2005) Arch Biochem Biophys 439:154–164

ANNEX 3

Kinetic and structural studies reveal a unique binding mode of sulfite to the nickel center in urease

Luca Mazzei [a], Michele Cianci [b], Stefano Benini [c], Leonardo Bertini [a], Francesco Musiani [a], Stefano Ciurli [a,*]

[a] Laboratory of Bioinorganic Chemistry, Department of Pharmacy and Biotechnology, University of Bologna, Via Giuseppe Fanin 40, I-40138, Bologna, Italy
[b] European Molecular Biology Laboratory c/o DESY, Notkestraße 85, 22607 Hamburg, Germany
[c] Bio-Organic Chemistry and Bio-Crystallography Laboratory, Faculty of Science and Technology, Free University of Bolzano, Piazza Università 5, 39100 Bolzano, Italy

ARTICLE INFO

ABSTRACT

Urease is the most efficient enzyme known to date, and catalyzes the hydrolysis of urea using two Ni(II) ions in the active site. Urease is a virulence factor in several human pathogens, while causing severe environmental and agronomic problems. Sporosarcina pasteurii urease has been used extensively in the structural characterization of the enzyme. Sodium sulfite has been widely used as a preservative in urease solutions to prevent oxygen-induced oxidation, but its role as an inhibitor has also been suggested. In the present study, isothermal titration microcalorimetry was used to establish sulfite as a competitive inhibitor for S. pasteurii urease, with an inhibition constant of 0.19 mM at pH 7. The structure of the urease–sulfite complex, determined at 1.65 Å resolution, shows the inhibitor bound to the dinuclear Ni(II) center of urease in a tridentate mode involving bonds between the two Ni(II) ions in the active site and all three oxygen atoms of the inhibitor, supporting the observed competitive inhibition kinetics. This coordination mode of sulfite has never been observed, either in proteins or in small molecule complexes, and could inspire synthetic coordination chemists as well as biochemists to develop urease inhibitors based on this chemical moiety.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Urease (urea aminohydrolase, EC 3.5.1.5) is a nickel-dependent non-redox enzyme whose catalytic function is the hydrolysis of urea to yield ammonia and carbamate at a rate 10^{15} times higher than the uncatalyzed reaction, making it the most efficient enzyme known to date [1]. Carbamate then spontaneously evolves to produce another molecule of ammonia and carbon dioxide. Urease catalyzes this last step of organic nitrogen mineralization in bacteria, fungi, plants, algae and invertebrates [2–5]. The overall hydrolysis of the products generated by urease activity determines an increase in pH of the surrounding milieu, causing negative consequences in medical and agricultural settings [2–7]. This enzyme represents the main virulence factor for a large variety of ureolytic human pathogens such as Helicobacter pylori [8], Mycobacterium tuberculosis [9], Yersinia enterocolitica [10], Cryptococcus neoformans [11], and Proteus mirabilis [12]. Furthermore, ureolytic bacteria expressing urease are widespread in soils that are treated with urea, a nitrogen fertilizer used worldwide [13], and their activity contributes to a number of significant environmental and economic problems such as loss of nitrogen from soil and release of ammonia in the atmosphere, ammonia toxicity for plants, and seedlings damage [14]. In all these instances, a tight control of urease activity is

required to counteract its deleterious effects. For this purpose, several classes of molecules have been proposed and tested, both in medicine and agriculture, as urease inhibitors [4,15–18].

Knowledge about the structure of the nickel-containing active site cavity has been derived through studies on native ureases isolated from several sources [19–22], which revealed a conservation in the coordination environment around the two Ni(II) ions (Fig. 1A).

On the basis of structural information obtained from the crystal structures of native Sporosarcina pasteurii (formerly known as Bacillus pasteurii) urease (SPU) [18,20] and of its complexes with a range of other ligands [18,20,23–27], a general scheme of the catalytic mechanism of ureases was proposed [1,5,20,28–31]. In this mechanism, the nickel-bridging hydroxide acts as the nucleophilic group that attacks the urea molecule chelating the bimetallic nickel cluster using an oxygen atom and a nitrogen atom. An extended network of second shell hydrogen bonds appears to stabilize the substrate binding during the catalytic process. In addition, a flexible flap changes the active site channel from an open to a closed conformation (Fig. 1B).

Sulfite has been extensively used as a preservative in solutions of jack bean (Canavalia ensiformis urease, JBU) [32,33] and SPU [18,20, 23–27,34,35]. The role of sulfite as a stabilizer of the urease activity has been interpreted in the past as due to the maintenance of the redox state of the conserved cysteine residue on the enzyme active site flap, which is essential for enzyme activity in its reduced thiol form (Fig. 1B, C). Indeed, it is known that the thiol groups of cysteines

* Corresponding author.
  E-mail address: stefano.ciurli@unibo.it (S. Ciurli).
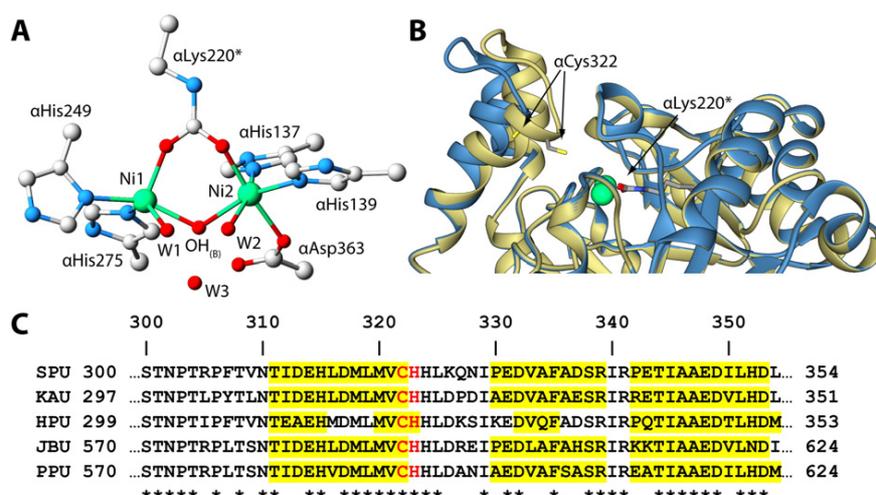
Fig. 1. (A) Details of the active site of native SPU. Atoms are colored according to the atom type. (B) Superimposition of the open (blue ribbons, 4CEU) and closed (yellow ribbons, 3UBP) conformation of the flexible flap. Ni(II) ions are shown as green spheres. The αCys322 side chain is reported as "sticks". (C) Multiple sequence alignment of the flap region of the ureases for which a crystal structure is available: SPU, *Klebsiella aerogenes* urease (KAU), *Helicobacter pylori* urease (HPU), jack bean urease (JBU), and pigeon pea urease (PPU). The position of SPU αCys322 and αHis323 are in red, while the α-helices are highlighted in yellow. The asterisks indicate the fully conserved amino acids.

can form oxygen derivatives such as sulfenic (Cys-SOH), sulfinic (Cys-SO$_2$H), and sulfonic (Cys-SO$_3$H) functionalities [36], while stable sulfenic acids may be produced by mild oxidation of sterically hindered thiols [37]. In the case of the sulfenic functional group, the role of sulfite could be the intermediate formation of an enzymatically inactive S-sulfocysteine (Cys-S-SO$_3^-$), in turn undergoing hydrolysis to sulfate and back to Cys-SH:

$$Cys{-}S{-}OH + HSO_3^{\ominus} \longrightarrow Cys{-}S{-}SO_3^- + H_2O \longrightarrow Cys{-}SH + HSO_4^{\ominus}$$
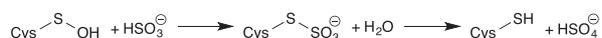
In addition, it has been shown that the enzymatic activity of JBU solutions, stored in the presence of β-mercaptoethanol (BME) but in contact with air oxygen, decreases [32] because of the formation of a mixed disulfide bond involving BME and the conserved active site flap Cys592 [22]. This process is reverted by treatment of inactivated urease with sulfite [32] possibly through the reduction of this disulfide bond [38]:

$$Cys{-}S{-}S{-}BME + HSO_3^{\ominus} \longrightarrow Cys{-}SH + {}^{\ominus}O_3S{-}S{-}BME$$

The opposite action of BME and sulfite on JBU was also suggested on the basis of an increase in the anodic electrophoretic mobility of native urease treated with sodium sulfite caused by an increased negative charge of the protein molecule; this process could be reversed by treatment with BME and was explained by the formation of S-sulfocysteines involving the several cysteine residues on the surface of JBU [39].

In addition to a protective action on the urease active site cysteine, sulfite has been reported to be an inhibitor of JBU, with a competitive inhibition mechanism deduced from the temperature dependence of the activation energy of the urea enzymatic hydrolysis as a function of sulfite concentration [40]. A pH-dependent kinetic study further suggested that the bisulfite mono-anion, and not the sulfite di-anion, is the actual inhibitor of JBU with inhibition constants in the milli-molar range, and that this process does not entail an interaction of bisulfite with the sulfhydryl group of essential cysteines, but rather an addition to the active site, whose nature was still very obscure at the time [41]. These early reports were later supported by a study that indicated sulfite acting as a competitive inhibitor for JBU with an inhibition constant K$_i$ = 2.23 ± 0.45 mM at pH 7.0 [42]. This inhibition role of sulfite on urease was then proposed to involve a direct interaction with the Ni(II) ions in the active site on the basis of an apparent increase of the nickel affinity of JBU in the presence of this anion [38].

In this study, we report a molecular characterization of the inhibition of SPU by sulfite. In particular, we describe and discuss the results of pH-dependent kinetic measurements carried out using a calorimetry-based assay, indicating that sulfite acts as a competitive inhibitor of SPU. This conclusion is supported by the crystal structure of the SPU–sulfite complex at 1.65 Å resolution, showing an unprecedented binding of the inhibitor to the active site Ni(II) ions through its three oxygen atoms.

2. Materials and methods

2.1. Protein purification

S. pasteurii DSM 33 cells were obtained using a modification of a previously described procedure [18], and SPU was isolated in a pure form to a specific activity of about 2500 units per milligram using the following protocol. *S. pasteurii* cells, resuspended in a buffer containing 50 mM phosphate pH 7.5, 50 mM Na$_2$SO$_3$, 1 mM EDTA (buffer A), additionally containing 10 mM MgCl$_2$ and 20 μg/mL DNAse I, were disrupted by three passages through a French® pressure cell press (SLM Aminco) at 20,000 psi. The soluble fraction was obtained by differential centrifugation, first at 30,000 ×g for 30 min and then at 150,000 ×g for 2 h, in order to remove cell debris and aggregates. The crude extract was dialyzed overnight against buffer A and then loaded onto a Q Sepharose XK 50/20 (GE Healthcare) anionic exchange column, previously equilibrated with the same buffer. A step gradient procedure was used to elute the protein with increasing ionic strength (NaCl in buffer A was used at concentrations of 150, 350, 450 mM) at a flow rate of 5 mL min^{-1}. The active fractions, eluted and detected at a concentration of 350 mM NaCl, were pooled and the ionic strength raised to 1 M using (NH$_4$)$_2$SO$_4$. After centrifugation at 30,000 ×g for 15 min to remove the precipitate, the solution was loaded onto a Phenyl Sepharose XK 26/20 (GE Healthcare) hydrophobic interaction column equilibrated with buffer A containing 1 M (NH$_4$)$_2$SO$_4$ and eluted with a linear gradient from 1 M to 0 M at a flow rate of 3 mL min^{-1}. Urease, eluted with 400 mM (NH$_4$)$_2$SO$_4$, was concentrated using an Amicon ultrafiltration cell equipped with a membrane of 100,000 Dalton molecular weight cut-off. The resulting enzyme solution was loaded onto a Superdex 200 XK 16/60 (GE Healthcare) gel filtration column equilibrated with buffer A, containing 150 mM NaCl in order to prevent non-specific

interaction between the protein and the resin, and eluted at a flow rate of 1 mL min^{-1}. The active fractions were pooled, and the ionic strength was decreased by dilution. The obtained solution was loaded onto a Mono Q HR 10/10 (Pharmacia) anionic exchange column, equilibrated with buffer A. Urease was eluted using a linear gradient of increasing ionic strength from 0 to 500 mM NaCl in buffer A. The active fractions were concentrated and further purified to homogeneity using a Superdex 200 10/300 GL (GE Healthcare) gel filtration column equilibrated with 20 mM N-(2-hydroxyethyl)piperazine-N′-ethane-sulfonic acid sodium salt (HEPES), pH 7.5, containing 150 mM NaCl and 50 mM Na$_2$SO$_3$ (buffer B) at a flow rate of 0.5 mL min^{-1}. Urease active fractions from the last step were concentrated to 11 mg/mL and stored in buffer A at 4 °C.

2.2. Kinetic studies

Calorimetric experiments were carried out with a high-sensitivity VP-ITC (ITC: Isothermal Titration Calorimetry) micro-calorimeter (MicroCal LLC, Northampton, MA, USA), using a previously described enzymatic assay [18,43]. SPU was eluted through a Superdex 200 10/300 GL (GE Healthcare) size-exclusion chromatography column immediately before the measurements, using a 50 mM HEPES pH 7.5 buffer, containing 50 mM Na$_2$SO$_3$, 2 mM EDTA, and 150 mM NaCl, as eluent. The protein samples were then diluted to the concentrations used for the calorimetric experiments described below, using 50 mM HEPES containing 2 mM EDTA, at the pH value (7.0, 7.5 or 8.0) used in the experiment. The obtained enzyme solution was degassed and loaded into the ITC sample cell (V = 1.4093 mL). The ITC stirring injection syringe was filled with urea dissolved in 50 mM HEPES containing 2 mM EDTA to the concentrations used for the different experiments described below. The reference cell was filled with deionized water, and the temperature of the two cells was set and stabilized at 298 K. Stirring speed was 300 rpm, and thermal power was monitored every 2 s using high instrumental feedback. In order to determine the molar enthalpy of the reaction (M1 experiments), 30 nM SPU and 20 mM urea were used in cell and in syringe, respectively, and a single injection of 5 μL urea was carried out, giving a final substrate concentration of 0.07 mM in the sample cell. After the baseline returned to the original level, indicating that the consumption of the substrate was complete, a second and a third injection were performed, following the same procedure. Numerical integration of the area under the three single peaks was carried out using the native Origin package provided by the calorimeter manufacturer, and the average value thus obtained was used to calculate the molar enthalpy for urea hydrolysis. In order to derive the kinetic parameters K_m and k_{cat} (M2 experiments) 50 pM SPU and 0.5 M urea were used in the cell and in the syringe, respectively. The enzyme was pre-incubated for 15 min at 25 °C in 50 mM HEPES, 2 mM EDTA, at pH 7.0, 7.5 or 8.0, and in the presence of a predefined amount of sulfite depending on the experiment (see below). Successive injections of 5 μL urea solution, containing sulfite at the same concentration of the enzyme solution, were carried out every 180 s, a time necessary to allow the thermal trace to return to a steady-state level after each injection and, at the same time, maintain pseudo-first-order reaction conditions. The thermal power obtained from the baseline shift was averaged using the last 15 s prior to the subsequent injection to obtain an accurate measurement. The calculated thermal power for each injection was converted to reaction rate and corrected to take into account enzyme dilution. Control runs were carried out by injecting the urea solution into the buffer alone. All calorimetric experiments were carried out using a minimum required amount of sulfite (20 μM), which was strictly necessary in order to keep the enzyme active during the course of the measurement. Indeed, analogous experiments performed in the total absence of sulfite indicated a progressive decrease of enzyme activity. Data were processed with the native Origin package. The reaction rates were obtained using a previously described protocol [18,43].

The data analysis was based on the classic definitions of competitive, uncompetitive and mixed enzyme inhibition [44,45]. In the most general case, the rate of enzymatic reaction is given by:

$$v = \frac{V_{max} \cdot [\text{urea}]}{K_m \cdot \left(1 + \frac{[\text{I}]}{K_{ic}}\right) + [\text{urea}] \cdot \left(1 + \frac{[\text{I}]}{K_{iu}}\right)}. \tag{1}$$

In this expression, $V_{max} = k_{cat}[\text{urease}]$ and K_m are the maximum rate and the Michaelis constant, respectively, [I] is the molar concentration of the inhibitor, and K_{ic} and K_{iu} are the competitive and uncompetitive inhibition constants, respectively. The kinetic parameters were obtained from a fit of the experimental reaction rates obtained by calorimetry to the general inhibition Eq. (1), using non-linear regression analysis implemented in the Origin package.

2.3. Crystallization, data collection and structural determination

Crystallization was performed at 293 K using the hanging-drop method. A solution of urease (11 mg mL^{-1}, 3 μL) in 20 mM HEPES, pH 7.5, containing 150 mM NaCl and 50 mM Na$_2$SO$_3$, was diluted with 3 μL of the precipitant solution (1.8–2.0 M ammonium sulfate in a 50 mM sodium citrate buffer pH 6.3 containing 50 mM Na$_2$SO$_3$ and 1–5% dimethylsulfoxide, DMSO). The drop was equilibrated by vapor diffusion against 1 mL of the precipitant solution using a Hampton Research 24-well Linbro plate. Rice-shaped protein crystals appeared within 1–3 weeks and grew to a size of 0.1 × 0.1 × 0.3 mm^3. Crystals were scooped up using cryoloops, transferred to a cryoprotectant solution containing 20% ethylene glycol, 2.4 M ammonium sulfate, 50 mM sodium citrate and 50 mM Na$_2$SO$_3$, and then flash-cooled and stored in liquid nitrogen.

Diffraction data were collected at 100 K using synchrotron radiation at the EMBL P13 beamline of the Petra III storage ring, c/o DESY Hamburg (Germany), equipped with a Si(111) crystal monochromator (FMB Oxford), a DECTRIS Pilatus 6 M detector, and a MD2 goniometer (MAATEL-EMBL) with a horizontal spindle axis. Reflection data were collected from two crystals (1500 images for each crystal), by performing helical scans through the crystal to achieve higher resolution and minimize radiation damage. The data were processed using XDS [46] and AIMLESS [47]. The crystals are isomorphous with those of apo-SPU and other complexes of this enzyme (Table 1).

The model of SPU in complex with citrate (PDB code 4AC7, 1.50 Å resolution) [27], devoid of water molecules and ligands, was used as a starting model for the rigid body refinement of the single αβγ protein trimer, carried out using Refmac [48,49]. Manual model building was conducted using Coot [50,51]. The structure was isotropically refined, including the hydrogen atoms in the riding positions. The diffraction data and final refinement statistics are given in Table 1. The structure was deposited in the Protein Data Bank under the accession code 5A6T. Crystallographic figures were created using Chimera [52] or CrystalMaker 8.7 (http://www.crystalmaker.com/).

2.4. Theoretical calculations

Density functional theory (DFT) computations were carried out using the program ORCA 3.0.3 [53] and the Becke three-parameter hybrid functional combined with Lee–Yang–Parr correlation functional (B3LYP/G) [54,55] as defined in the Gaussian software [56]. All atoms were described by the Dunning correlation-consistent polarized triple zeta basis set with the inclusion of diffuse functions (aug-cc-pVTZ) [57]. Frequency computations were performed to determine the nature of the various critical points.

Table 1
X-ray diffraction data collection and refinement statistics for sulfite-bound *Sporosarcina pasteurii* urease (PDB code 5A6T).

Data collection	
Wavelength (Å)[a]	1.033
Crystal-to-Detector distance (mm)	271.8
Oscillation angle (degrees)	0.075
Number of images	3000
Space group	$P6_322$
Unit cell (*a*, *b*, *c*, Å)	131.24, 131.24, 188.88
Resolution range (Å)[a]	97.46–1.65 (1.68–1.65)
Total number of reflections[a]	2,478,074 (69,467)
Unique reflections[a]	115,020 (5576)
Multiplicity[a]	21.5 (12.5)
Completeness[a] (%)	99.8 (99.2)
R_{sym}[a,b] (%)	0.102 (0.471)
R_{pim}[a,c] (%)	0.030 (0.199)
Mean I half-set correlation CC(1/2)	0.969
Mean $I/\sigma(I)$[a]	29.4 (5.7)
Refinement statistics	
Number of monomers in the asymmetric unit	3
R_{factor}[d] (%)	12.87
R_{free}[d] (%)	15.25
Cruickshank's DPI for coordinate error[e] based on R_{factor} (Å)	0.063
Wilson plot B-factor (Å²)	12.9
Average all atom B-factor[f] (Å²)	19.155
B-factor of Ni atoms	Ni1: 18.5; Ni2: 16.1
B-factor of sulfite atoms	S: 25.1, O1: 23.5; O1: 28.2; O1: 22.1
RMS (bonds)[d]	0.022
RMS (angles)[d]	2.067
Total number of atoms[c]	6864
Total number of water molecules	636
Solvent content (%)	55.23
Matthews coefficient (Å³/Da)	2.75
Ramachandran plot[g]	
Most favored region (%)	90.7
Additionally allowed region (%)	8.4
Generously allowed region (%)	0.8
Disallowed region (%)	0.2

[a] Highest resolution bin in parentheses.
[b] $R_{sym} = \sum_{hkl} \sum_j |I_j - \langle I \rangle| / \sum_{hkl} \sum_j I_j$, where I is the intensity of a reflection, and $\langle I \rangle$ is the mean intensity of all symmetry related reflections j.
[c] $R_{p.i.m.} = \sum_{hkl} \{[1/(N-1)] \sum_j |I_j - \langle I \rangle|\}^{1/2} / \sum_{hkl} \sum_j I_j$, where I is the intensity of a reflection, and $\langle I \rangle$ is the mean intensity of all symmetry related reflections j, and N is the multiplicity [76].
[d] Taken from REFMAC [48,49]; R_{free} is calculated using 5% of the total reflections that were randomly selected and excluded from refinement.
[e] $DPI = R_{factor} \cdot D_{max} \cdot compl^{-1/3} \sqrt{\frac{N_{atoms}}{(N_{refl} - N_{params})}}$, where N_{atoms} is the number of the atoms included in the refinement, N_{refl} is the number of the reflections included in the refinement, D_{max} is the maximum resolution of reflections included in the refinement, *compl* is the completeness of the observed data, and for isotropic refinement, $N_{params} \approx 4N_{atoms}$ [77].
[f] Taken from BAVERAGE [78].
[g] Taken from PROCHECK [78].

3. Results

3.1. Calorimetric analysis

The inhibition of *S. pasteurii* urease by sulfite was investigated using isothermal titration calorimetry (ITC), a method that provides a universal approach to determining the kinetic behavior of enzymes by yielding, in a single experiment, a complete set of kinetic parameters for an enzyme-catalyzed reaction [43,58–61]. A very similar approach has been recently described for other metallo-hydrolases [62].

The molar reaction enthalpy (ΔH) of the hydrolysis of urea by urease was determined using a so-called M1 experiment, in which a single injection of the substrate solution into a concentrated enzyme solution determined a decrease of the instrumental thermal power necessary to maintain the reference and the sample cell at a constant temperature, indicating an exothermic reaction. Complete consumption of the substrate occurred in the range 500–2000 s, depending on the working

pH in the 7.0–8.0 range (Fig. 2A). Two further injections of substrate into the reaction cell provided curves with identical profile, showing negligible inhibition by products (Fig. 2A). The integration of the curves yielded $\Delta H = -11.3 \pm 0.2$ kcal mol^{-1}. A progressive increase of the reaction rate with increasing pH was observed in the 7.0–8.0 range, consistent with the known pH dependence of the enzyme activity for SPU [63].

The so-called M2 experiments (Fig. 2B) were performed using a diluted enzyme solution in the sample cell and carrying out multiple injections of a concentrated substrate solution, in the pH range 7.0–8.0. These experiments revealed an initial increase in thermal power due to the heat of substrate dilution, followed by a decrease required to maintain isothermal conditions for the exothermic reaction. The rate of heat generated by the enzyme is equivalent to the decrease in thermal power after each injection, steadily increasing as the substrate concentration increases. Analogous experiments carried out using 1.6 mM sulfite (Fig. 2B) showed a decrease of the reaction rate. Monitoring the latter as a function of substrate concentration and pH (in the range 7.0–8.0), as shown in Fig. S1 (Supplementary information), yielded a series of data that were fit to Eq. (1), resulting in the kinetic parameters reported in Table 2. The increase of k_{cat} as a function of pH in the explored range is in complete agreement with previous literature data indicating an optimum pH for SPU of ca. 7.7 [63]. The Michaelis–Menten constant, slightly smaller than the value of 17.3 mM reported in phosphate buffer [34,63], is invariant with pH. The parameters in Table 2 further show values of $K_{ic} \ll K_{iu}$, indicative of a competitive mechanism of inhibition of urease by sulfite in the investigated pH range. This inhibition effect strongly decreases upon pH increase, becoming undetectable at pH 8.0.
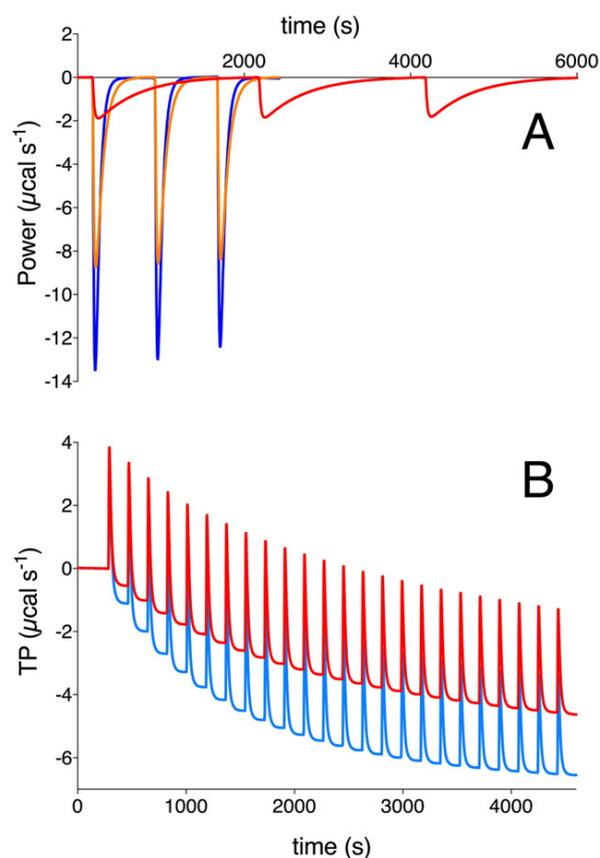
Fig. 2. Typical plots of thermal power of *S. pasteurii* urease activity as a function of time in M1 experiments (A) at pH 7.0 (red line), pH 7.5 (orange line) and pH 8.0 (blue line), and in M2 experiments at pH 7.5 in the presence of 0.02 mM (blue line) and 1.6 mM (red line) sodium sulfite (B). Experimental conditions are described in Materials and Methods.

Table 2
Kinetic parameters for the inhibition of SPU with sulfite.

pH	k_{cat} (s^{-1})	K_M (mM)	K_{ic} (mM)	K_{iu} (mM)
7.0	7953 ± 36	10.3 ± 0.1	0.19 ± 0.03	»1
7.5	11,631 ± 70	12.0 ± 0.2	1.17 ± 0.02	»1
8.0	9939 ± 49	11.7 ± 0.2	–	–

3.2. Structural analysis

The crystal structure of sulfite-inhibited SPU, showing the well-established architecture of this metallo-enzyme, closely matches that of the native protein (PDB code 4CEU, used here as reference for comparison because it has been refined to the highest resolution so far available, 1.58 Å) [18]. The global pairwise root mean square deviation (RMSD) per residue between Cα atoms of the two structures is only 0.11, 0.06 and 0.05 Å for the α, β and γ subunits, respectively, and the overall RMSD is 0.10 Å. The analysis of the residue-averaged backbone B factors, as well as the φ and φ protein backbone angles (Figs. S2–S4), reveals two regions in the α subunit that feature a significantly larger mobility than the rest of the protein. Region A (residues 380–405) is located on a surface patch and involves a helix and a solvent-exposed loop. This region has been proposed [27] to participate in either the catalytic mechanism or the interaction with UreD, the accessory protein that, along with its partners UreF, UreG and UreE, forms the supercomplex that is necessary for the delivery of the two Ni(II) ions into the active site cavity and for the consequent urease activation [5,64]. Region B (residues 305–350) corresponds to the flap, consisting of a helix–loop–helix motif, responsible for the modulation of the substrate access to the active site cavity (Fig. 1B). The sulfite-bound SPU structure displays an "open" flap, as observed in the case of native urease [18,20], as well as in all the other cases of SPU-inhibitor complex structures so far determined [18,23–27], except for urease inhibited by diamidophosphate (DAP), whose crystal structure presents the flap in the completely "closed" position [20]. Since DAP is generated in situ from the hydrolysis of phenylphosphorodiamidate, simulating the catalytic intermediate or transition state, the closed-flap conformation was proposed to stabilize the intermediate or transition state of the catalysis [20]. In the refined model (Figs. 3A and 4), the two Ni(II) ions are bridged by the carboxylate group of the carbamylated Lys220*, which is bound to Ni(1) by Oθ1 and to Ni(2) by Oθ2, allowing the two Ni(II) ions to be properly placed for the catalytic process. Ni(1) is further coordinated by His249 Nδ and by His275 Nε, whereas Ni(2) is bound to His137 Nε, His139 Nε, and Asp363 Oδ1.

The unbiased electron density map calculated with Fourier coefficients F_o–F_c and phases derived from the model of SPU in complex with citrate (PDB code 4AC7, 1.50 Å resolution) [27], devoid of water molecules and ligands, displays an electron density in the volume around and between the Ni(II) ions that clearly indicates the existence of a trigonal pyramidal non-protein ligand bound to the two metal ions (Fig. 3B). This electron density is very different from what would be expected if four solvent-derived molecules were present in the active site (see, for example, the maps reported in [18]). Given the presence of sulfite in the crystallization solution, this moiety was modeled into the electron density and refined with full occupancy (Fig. 4). In the refined model, the sulfite molecule is bound in a tridentate binding mode to the two Ni(II) ions in the active site, replacing the terminal and the bridging solvent molecules that were bound to Ni(1) and Ni(2) in the native form of the enzyme. This type of coordination is unprecedented both in protein and in small molecules structures. In this mode, sulfite binds the two Ni(II) ions at the active site using two oxygen atoms [Ni(1)–O(1) and Ni(2)–O(2), respectively], placing its third oxygen atom, O(B), in the bridging position. The positions of the conserved amino acid residues, which are not involved in the Ni(II) binding but are known to be crucial in the catalytic mechanism (Ala170, His222, Glu223, Asp224, Gly280, His323, Ala366 and Met367), are largely
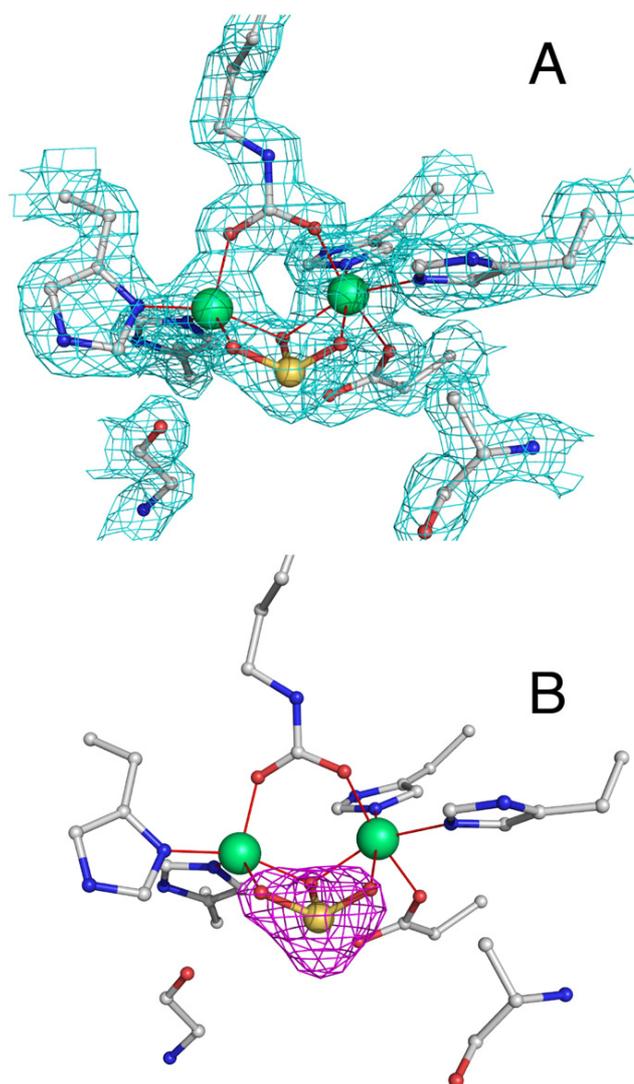
Fig. 3. Atomic model of the active site of sulfite-bound *Sporosarcina pasteurii* urease (PDB code 5A6T). In panel (A), the nickel-coordination environment is shown superimposed on the final 2F_o–F_c electron density map contoured at 1.5 σ (cyan). In panel (B), the unbiased electron density map calculated with Fourier coefficients F_o–F_c and phases derived from the model before the atoms corresponding to S, O(1), O(2), and O(B) of sulfite were added, is shown contoured at 3 σ (magenta). The carbon, nitrogen, oxygen, sulfur and nickel atoms are gray, blue, red, yellow and green, respectively.

conserved between the native and inhibited urease. This conservation of the structure of the active site residues is generally observed also in all other structures of SPU.

A comparison of the geometric parameters around the two Ni(II) ions in the active site of native and sulfite-bound urease, presented in Table S1, reveals that the structure is rigidly maintained.

The molecule of sulfite bound to urease is slightly asymmetric, with S–O(1) = 1.67 Å, S–O(2) = 1.57 Å, and S–O(B) = 1.83 Å. Moreover, the sulfite molecule forms a series of hydrogen bonds with the residues that surround the active site cavity (Fig. 4): O(1) forms a hydrogen bond with His222 Nε (at 2.65 Å) that is protonated and acts as a hydrogen-bond donor, as deduced from the interaction of His222 Nδ with the Asp224 peptide NH group at 2.88 Å. Furthermore, two water molecules are H-bonded to O(1) at 3.19 Å and 3.21 Å. The sulfite oxygen O(2), bound to Ni(2), is placed at 2.78 Å from the carbonyl backbone Ala170 O, suggesting a possible H-bond in which the sulfite O(2) could be protonated. Finally, the bridging sulfite O(B) is placed at 2.48 Å from
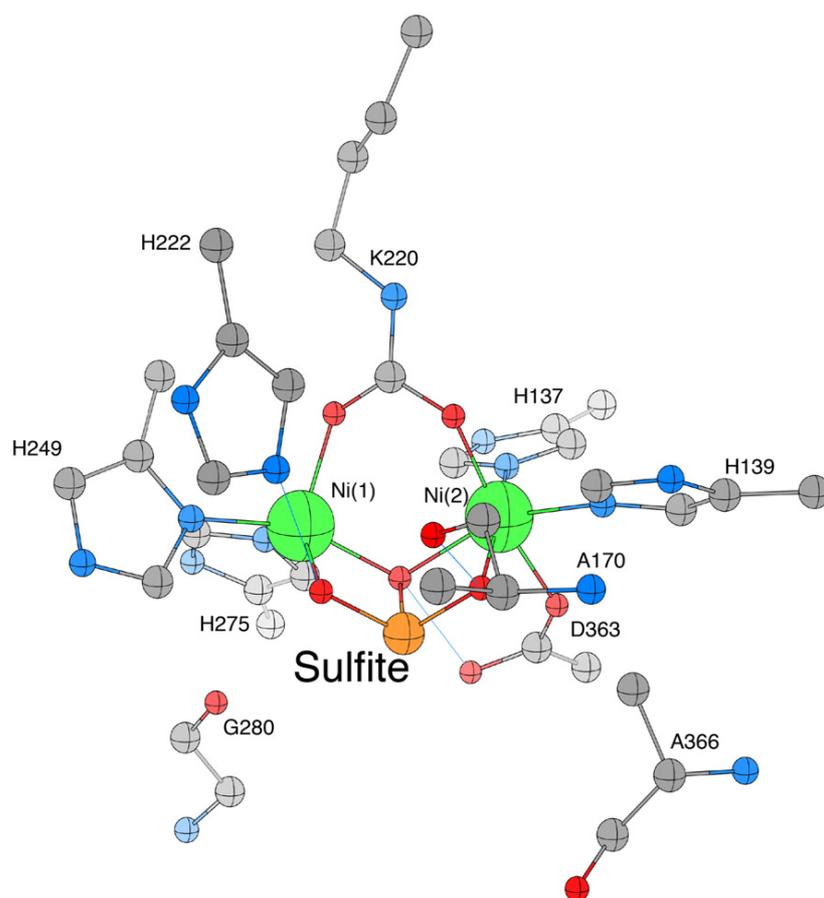
Fig. 4. Crystallographic structural model for the active site obtained for the sulfite-bound *Sporosarcina pasteurii* urease (PDB code 5A6T). The carbon, nitrogen, oxygen, sulfur and nickel atoms are gray, blue, red, yellow and green, respectively. Putative H-bonds are shown as thin blue lines. Spheres are drawn using the relative atomic radii values in CrystalMaker.

Asp363 Oδ2, again suggesting that these two atoms could share a hydrogen atom through an H-bond.

4. Discussion

The present study focuses on the biochemical and structural features of the interaction of the urease from *S. pasteurii* with sulfite, a known urease inhibitor [40–42] as well as a stabilizer of urease activity [18, 20,23–27,32–35]. Enzymatic measurements of sulfite inhibition of urease were performed on SPU using a calorimetric method, and the results were integrated with the analysis of high-resolution crystallographic data of sulfite-inhibited urease. The calorimetric data show that sulfite inhibits the activity of SPU with a pH-dependent competitive mechanism in the 7.0–8.0 range. These results, obtained on SPU, are in agreement with early data on JBU [41], and suggest a common mechanism of urease inhibition by sulfite among ureases from different sources. Among the several classes of urease inhibitors known to date [4], the competitive inhibition constant determined for sulfite at pH 7.0 (0.19 mM) compares well with the analogous value reported for boric acid (0.08–0.35 mM) and BME (0.72 mM), while it performs better then urea analogs (1.45–980 mM), phosphate (19 mM), and bismuth derivatives (1.74–1.84 mM).

The competitive nature of the inhibition indicates that sulfite, in one of its protonated forms, competes with the substrate urea for the active site Ni(II) ions. In particular, the inhibition decreases almost an order of magnitude by increasing the pH from 7.0 to 7.5, and becomes negligible upon further increase of the pH up to 8: this observation excludes that the inhibitor takes the di-anionic SO_3^{2-} form, but does not in principle distinguish between the mono-anionic and the neutral non-dissociated form

of sulfite ($pK_{a1} = 1.85$ and $pK_{a2} = 7.20$ [65]) because of the limited pH interval that could be explored without causing enzyme degradation at low pH. Even though the crystallization buffer pH = 6.3 would support the presence of the mono-anionic inhibitor as the main form of sulfite in solution, the presence of the neutral form of the inhibitor in the urease active site cannot be excluded on the sole basis of these criteria.

Consistent with these kinetic observations, high resolution crystallographic data demonstrate that the inhibitor binds in the active site displacing three solvent-derived Ni(II)-bound molecules and replacing these position with its three oxygen atoms, blocking access of urea to the key metal ions. This tripodal coordination mode of sulfite has never been observed either in proteins or in small molecules complexes, and represents an exciting twist in the chemistry of Ni(II) in urease. Binuclear metallo-hydrolases are not new to this kind of tripodal ligand coordination mode: sulfate binds the Mn-Mn center of bacteriophage λ protein phosphatase [66], while phosphate binds the Fe-Mn center in purple acid phosphatase from sweet potato [67] and *Lupinus luteus* [68], as well as the di-nickel center of urease [25], using three out of the four oxygen atoms of the anion.

The structure of the bisulfite ion has actually been the subject of some debate: the HSO_3^- ion could adopt either a C_{3v} structure (HSO_3^-), in which the H atom is bound to sulfur, or a C_1 structure $[SO_2(OH)]^-$, in which a OH group is bound to S. Previous experimental data and quantum mechanical calculations showed that, in the first case, the three S–O bonds should be identical (bond distance ca. 1.47 Å, bond order = 5/3), while in the second case the S–OH bond would be longer (ca. 1.75–1.78 Å, bond order = 1) than the other two S–O bonds (ca. 1.49–1.50 Å, bond order = 3/2) [69]. In aqueous solution the C_1 isomer is more energetically stable than the C_{3v} isomer by only 4.9 kcal mol^{-1}

Table 3

B3LYP/G aug-cc-pVTZ distances (Å) of sulfurous acid, bisulfite ion in the C_1 and in the C_{3v} forms, and sulfite ion compared with the distances found in the sulfite inhibited SPU crystal structure.

	H_2SO_3	HSO_3^- (C_1)	HSO_3^- (C_{3v})	SO_3^{2-}	Crystal structure
$S-O_1$	1.647	1.782	1.440	1.554	1.87
$S-O_2$	1.647	1.492	1.440	1.554	1.57
$S-O_3$	1.469	1.506	1.440	1.554	1.67
O_1-H_1	0.972	0.967			
O_2-H_2	0.972				
$S-H_S$			1.343		

[70], suggesting that the adopted structure would depend on the chemical environment [69]. The sulfite ion SO_3^{3-}, on the other hand, can adopt only C_{3v} geometry with identical S–O bond distances of 1.53 Å in solution (bond order = 4/3) [71]. In order to distinguish between the different protonated forms of sulfite, we carried out quantum mechanical calculations, and compared the computed structural parameters with those determined in the crystal structure of sulfite-inhibited SPU (Tables 3 and S2).

Our calculations largely confirm these previously reported values, additionally indicating that the neutral form H_2SO_3 would possess one short and two long S–O distances. The asymmetric collection of S–O bond distances found in the crystal structure of SPU bound to this inhibitor (1.57 Å, 1.67 Å, and 1.83 Å for S–O(1), S–O(2) and S–O(B) respectively), in a short–short–long distribution (Table 3), are neither compatible with the SO_3^{2-} ion (as also suggested by the previous considerations on the crystallization pH vs. the pK_a of HSO_3^-), nor with the C_{3v} isomer of this anion, nor with the neutral form of the inhibitor, supporting instead the presence of the $[SO_2(OH)]^-$ mono-anionic form of sulfite bound to the active site Ni(II) ions. In particular, the long S–O(B) distance and the short O(B)–Oδ2 Asp363 distance (2.48 Å) suggest the presence of a hydrogen shared between these two O atoms. On the other hand, the longer distance involving S–O(1) vs. S–O(2) could be determined by the presence of the hydrogen-bonding donor His222 NεH in the close vicinity of O(1) (at 2.65 Å) that could have an effect of decreasing the negative charge on O(1); this consideration would also justify the longer distance (2.20 Å) between Ni(1) and O(1) as compared to the distance (2.10 Å) between Ni(2) and O(2). In general, the S–O distances found in the crystal structure of the complex are ca. 0.1–0.2 Å longer than those expected for the bisulfite C_1 isomer in aqueous solution: this could be due to the presence of the two Ni(II) ions bound to the anion. Indeed, a similar effect has been observed in the case of the structure of human carbonic anhydrase II complexed with bisulfite (PDB code 2CBD) [72]: this structure, solved at 1.67 Å resolution, together with an older structure of the same enzyme solved at 2.0 Å resolution (PDB code 5CAC) [73], is the only case that shows a sulfite ion bound to a metal ion in a protein, namely Zn(II), in a terminal mode, using a presumably protonated O atom, as manifested by the network of H-bonds around it [72]. In this case, the length of the S–O(H) bond involving the oxygen atom bound to Zn(II) is 1.82 Å, a value essentially identical to the S–O(B) distance in the SPU–sulfite complex, while the other two S–O bond lengths, involving O atoms not bound to other metal ions, are 1.46–1.47 Å, consistent with the distances observed for the C_1 isomer of HSO_3^-. All these considerations further support the presence of the mono-anionic C1 form of sulfite. However, the possibility of an additional H-bond formed between sulfite O(2) and the carbonyl backbone Ala170 O does not rule out the presence of the neutral, fully protonated, sulfurous acid.

The presence of sulfite bound to metal ions in proteins is not unique: a research performed using the MetalPDB search server (http://metalweb.cerm.unifi.it) [74] reveals that, in addition to the above mentioned carbonic anhydrase II, sulfite is detected bound to the heme Fe atom in the active site of sulfite and nitrate reductases. In all those instances, however, sulfite is bound to the metal ion through the S atom and not using the whole set of O atoms as found in the active site of sulfite-inhibited *S. pasteurii* urease. A search of the Cambridge Crystallographic Database (http://www.ccdc.cam.ac.uk), containing small molecules, gave no hits for this coordination mode of sulfite, further reinforcing the novelty of this coordination mode. In this context, it is worth mentioning that the effect of sulfite on a newly discovered Ni-containing enzyme, lactate racemase (Lar), has been recently reported [75]. This enzyme contains an organometallic nickel-containing prosthetic group, and sulfite was shown to act as a potent mixed inhibitor as well as being able to stabilize Lar activity and delay Ni loss; however, the molecular details of these effects have not been clarified [75].

5. Conclusions

The kinetics of inhibition of *S. pasteurii* urease with sulfite, determined using an enzyme assay based on microcalorimetry, indicate that this molecule acts as a competitive inhibitor. This result is confirmed by the structure of the urease–sulfite complex, which shows the inhibitor bound to the active site Ni(II) ions and blocking access to urea, the natural substrate for this enzyme. The structure of the nickel-bound sulfite described in the present article is unprecedented and expands the coordination chemistry of this inorganic molecule to include a tridentate mode. The latter might inspire model inorganic chemists towards the synthesis of analogous complexes and investigate their chemical and reactivity properties. At the same time, the results illustrated here could stimulate research towards urease inhibitors based on oxidized sulfur functional groups, demonstrated here to be able to act as efficient modulators of enzyme activity.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.jinorgbio.2015.11.003. These data include MOL files and InChiKeys of the most important compounds described in this article.

References

[1] M.J. Maroney, S. Ciurli, Chem. Rev. 114 (2014) 4206–4228.
[2] H.L. Mobley, R.P. Hausinger, Microbiol. Rev. 53 (1989) 85–108.
[3] P.A. Karplus, M.A. Pearson, R.P. Hausinger, Acc. Chem. Res. 30 (1997) 330–337.
[4] B. Krajewska, J. Mol. Catal. B Enzym. 59 (2009) 9–21.
[5] B. Zambelli, F. Musiani, S. Benini, S. Ciurli, Acc. Chem. Res. 44 (2011) 520–530.
[6] B.P. Callahan, Y. Yuan, R. Wolfenden, J. Am. Chem. Soc. 127 (2005) 10828–10829.
[7] C. Follmer, Phytochemistry 69 (2008) 18–28.
[8] C. Montecucco, R. Rappuoli, Nat. Rev. Mol. Cell Biol. 2 (2001) 457–466.
[9] W. Lin, V. Mathys, E.L. Ang, V.H. Koh, J.M. Martinez Gomez, M.L. Ang, S.Z. Zainul Rahim, M.P. Tan, K. Pethe, S. Alonso, Infect. Immun. 80 (2012) 2771–2779.
[10] G.M. Young, D. Amid, V.L. Miller, J. Bacteriol. 178 (1996) 6487–6495.
[11] G.M. Cox, J. Mukherjee, G.T. Cole, A. Casadevall, J.R. Perfect, Infect. Immun. 68 (2000) 443–448.
[12] B.D. Jones, C.V. Lockatell, D.E. Johnson, J.W. Warren, H.L. Mobley, Infect. Immun. 58 (1990) 1120–1123.
[13] P.J.E.H.R.D. Stangel, American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, 1984 23–54.
[14] J.M.M. Bremner, R.L. Mulvaney, in: R.G. Burns (Ed.), Soil Enzymes, 1978.
[15] B. Krajewska, W. Zaborska, M. Leszko, J. Mol. Catal. B Enzym. 14 (2001) 101–109.
[16] W. Zaborska, B. Krajewska, Z. Olech, J. Enzyme Inhib. Med. Chem. 19 (2004) 65–69.
[17] P. Kosikowska, L. Berlicki, Expert Opin. Ther. Pat. 21 (2011) 945–957.
[18] S. Benini, M. Cianci, L. Mazzei, S. Ciurli, J. Biol. Inorg. Chem. 19 (2014) 1243–1261.
[19] E. Jabri, M.B. Carr, R.P. Hausinger, P.A. Karplus, Science 268 (1995) 998–1004.
[20] S. Benini, W.R. Rypniewski, K.S. Wilson, S. Miletti, S. Ciurli, S. Mangani, Structure 7 (1999) 205–216.
[21] N.C. Ha, S.T. Oh, J.Y. Sung, K.A. Cha, M.H. Lee, B.H. Oh, Nat. Struct. Biol. 8 (2001) 505–509.
[22] A. Balasubramanian, K. Ponnuraj, J. Mol. Biol. 400 (2010) 274–283.
[23] S. Benini, W.R. Rypniewski, K.S. Wilson, S. Ciurli, S. Mangani, J. Biol. Inorg. Chem. 3 (1998) 268–273.
[24] S. Benini, W.R. Rypniewski, K.S. Wilson, S. Miletti, S. Ciurli, S. Mangani, J. Biol. Inorg. Chem. 5 (2000) 110–118.
[25] S. Benini, W.R. Rypniewski, K.S. Wilson, S. Ciurli, S. Mangani, J. Biol. Inorg. Chem. 6 (2001) 778–790.
[26] S. Benini, W.R. Rypniewski, K.S. Wilson, S. Mangani, S. Ciurli, J. Am. Chem. Soc. 126 (2004) 3714–3715.
[27] S. Benini, P. Kosikowska, M. Cianci, L. Mazzei, A.G. Vara, L. Berlicki, S. Ciurli, J. Biol. Inorg. Chem. 18 (2013) 391–399.
[28] S. Ciurli, S. Benini, W.R. Rypniewski, K.S. Wilson, S. Miletti, S. Mangani, Coord. Chem. Rev. 190–192 (1999) 331–355.
[29] F. Musiani, E. Arnofi, R. Casadio, S. Ciurli, J. Biol. Inorg. Chem. 6 (2001) 300–314.
[30] S. Ciurli, Nickel and its surprising impact in nature, John Wiley & Sons, Ltd., 2007 241–277.
[31] S. Benini, F. Musiani, S. Ciurli, in: R. Kretsinger, V. Uversky, E. Permyakov (Eds.), Encyclopedia of metalloproteins, Springer, New York 2013, pp. 2287–2292.
[32] P.W. Riddles, R.K. Andrews, R.L. Blakely, B. Zerner, Biochim. Biophys. Acta Protein Struct. Mol. 743 (1983) 115–120.
[33] B. Zerner, Bioorg. Chem. 19 (1991) 116–131.
[34] S. Benini, C. Gessa, S. Ciurli, Soil Biol. Biochem. 28 (1996) 819–821.
[35] S. Ciurli, C. Marzadori, S. Benini, S. Deiana, C. Gessa, Soil Biol. Biochem. 28 (1996) 811–817.
[36] K.G. Reddie, K.S. Carroll, Curr. Opin. Chem. Biol. 12 (2008) 746–754.
[37] W.S. Allison, Acc. Chem. Res. 9 (1976) 293–299.
[38] G.J. King, B. Zerner, Inorg. Chim. Acta 255 (1997) 381–388.
[39] J.J. Martín de Llano, J.G. Gavilanes, Electrophoresis 13 (1992) 300–304.
[40] G.B. Kistiakowsky, R. Lumry, J. Am. Chem. Soc. 71 (1949) 2006–2013.
[41] J.F. Ambrose, G.B. Kistiakowsky, A.G. Kridl, J. Am. Chem. Soc. 72 (1950) 317–321.
[42] R.K. Andrews, School of Molecular and Microbial Sciences(Ph.D.) University of Queensland, Brisbane, Australia, 1986 (vol.).
[43] L. Mazzei, S. Ciurli, B. Zambelli, J. Vis. Exp. (2014), e51487.
[44] I.H. Segel, Enzyme kinetics: behavior and analysis of rapid equilibrium and steady-state enzyme systems, Wiley-Interscience, 1993.
[45] A. Cornish-Bowden, Perspect. Sci. 1 (2014) 74–87.
[46] W. Kabsch, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 125–132.
[47] P. Evans, Acta Crystallogr. D Biol. Crystallogr. 62 (2006) 72–82.
[48] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Acta Crystallogr. D Biol. Crystallogr. 53 (1997) 240–255.
[49] G.N. Murshudov, A.A. Vagin, A. Lebedev, K.S. Wilson, E.J. Dodson, Acta Crystallogr. D Biol. Crystallogr. 55 (1999) 247–255.
[50] P. Emsley, K. Cowtan, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2126–2132.
[51] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 486–501.
[52] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, J. Comput. Chem. 25 (2004) 1605–1612.
[53] F. Neese, Wiley Interdiscip. Rev.: Comput. Mol. Sci. 2 (2012) 73–78.
[54] A.D. Becke, J. Chem. Phys. 98 (1993) 1372–1377.
[55] C. Lee, W. Yang, R.G. Parr, Phys. Rev. B 37 (1988) 785–789.
[56] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian, Inc., Wallingford, CT, USA, 2009.
[57] T.H. Dunning, J. Chem. Phys. 90 (1989) 1007–1023.
[58] C. Spink, I. Wadso, Methods Biochem. Anal. 23 (1976) 1–159.
[59] B.A. Williams, E.J. Toone, J. Org. Chem. 58 (1993) 3507–3510.
[60] M.J. Todd, J. Gomez, Anal. Biochem. 296 (2001) 179–187.
[61] M.W. Freyer, E.A. Lewis, Methods Cell Biol. 79 (113) (2008).
[62] M.M. Pedroso, F. Ely, T. Lonhienne, L.R. Gahan, D.L. Ollis, L.W. Guddat, G. Schenk, J. Biol. Inorg. Chem. 19 (2014) 389–398.
[63] C. Marzadori, S. Miletti, C. Gessa, S. Ciurli, Soil Biol. Biochem. 30 (1998) 1485–1490.
[64] S. Quiroz-Valenzuela, S.C. Sukuru, R.P. Hausinger, L.A. Kuhn, W.T. Heller, Arch. Biochem. Biophys. 480 (2008) 51–57.
[65] D.D. Perrin, Ionisation constants of inorganic acids and bases in aqueous solution, Pergamon, Oxford, 1982.
[66] W.C. Voegtli, D.J. White, N.J. Reiter, F. Rusnak, A.C. Rosenzweig, Biochemistry 39 (2000) 15365–15374.
[67] G. Schenk, L.R. Gahan, L.E. Carrington, N. Mitic, M. Valizadeh, S.E. Hamilton, J. de Jersey, L.W. Guddat, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 273–278.
[68] S.V. Antonyuk, M. Olczak, T. Olczak, J. Ciuraszkiewicz, R.W. Strange, IUCr J. 1 (2014) 101–109.
[69] M.A. Vincent, I.J. Palmer, I.H. Hillier, J. Mol. Struct. THEOCHEM 394 (1997) 1–9.
[70] D.A. Horner, R.E. Connick, Inorg. Chem. 25 (1986) 2414–2417.
[71] L. Eklund, T.S. Hofer, A.B. Pribil, B.M. Rode, I. Persson, Dalton Trans. 41 (2012) 5209–5216.
[72] K. Håkansson, M. Carlsson, L.A. Svensson, A. Liljas, J. Mol. Biol. 227 (1992) 1192–1204.
[73] A.E. Eriksson, T.A. Jones, A. Liljas, Proteins 4 (1988) 274–282.
[74] C. Andreini, G. Cavallaro, S. Lorenzini, A. Rosato, Nucleic Acids Res. 41 (2012) D312–D319.
[75] B. Desguin, T. Zhang, P. Soumillion, P. Hols, J. Hu, R.P. Hausinger, Science 349 (2015) 66–69.
[76] M.S. Weiss, J. Appl. Crystallogr. 34 (2001) 130–135.
[77] D.W.J. Cruickshank, Acta Crystallogr. D Biol. Crystallogr. 55 (1999) 583–601.
[78] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G.W. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 235–242.

ANNEX 4

Reprinted from:

COMMUNICATION

CrossMark
click for updates

Inactivation of urease by 1,4-benzoquinone: chemistry at the protein surface†

L. Mazzei,[a] M. Cianci,[b] F. Musiani[a] and S. Ciurli*[a]

The high activity of urease, a Ni(II) enzyme, has several adverse effects on human health and agriculture, and its modulation needs the use of inhibitors. 1,4-Benzoquinone (BQ) irreversibly inactivates *Sporosarcina pasteurii* urease (SPU), with first order kinetics for both the inhibitor and the enzyme. This reaction is stoichiometrically quenched in the presence of sulphite. The 2.07 Å crystal structure of SPU bound to BQ shows the presence of a 1,4-hydroquinone moiety covalently bound to the thiol group of αCys322, a key residue found on the mobile flap regulating the substrate access to the active site. The 1.75 Å crystal structure obtained when sulphite is added to a solution of SPU previously incubated with BQ shows the presence of a 2,5-dihydroxy-benzenesulphonate moiety bound to the αCys322 thiol group. These data reveal how the active site cysteine reacts with a prototypical BQ moiety, found in a large number of natural substances potentially suitable to control the urease activity.

The rapid hydrolysis of urea catalysed by the Ni(II)-containing enzyme urease[1] in bacteria, fungi, algae and plants causes an abrupt increase of pH and, consequently, negative side effects on human health[2] and the environment.[3] Ureases show a quaternary structure made of a functionally minimal trimer, differently assembled in higher order structures depending on the biological source.[1] Each unit of the trimer hosts one enzyme active site, constituted by two Ni(II) ions (Ni1 and Ni2) bridged by a carbamylated lysine residue (αLys220* in SPU) and by a hydroxide ion (Fig. 1A). The Ni(II) ions are further coordinated by two histidine residues, and are distinguished by the presence of a carboxylate O atom from an aspartate residue bound to Ni2. The coordination geometry of each Ni(II) ion is completed by a
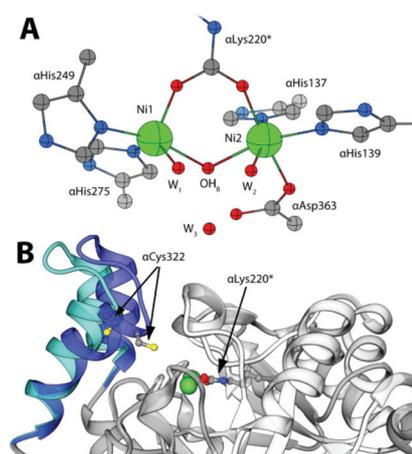
Fig. 1 (A) Coordination environment of the Ni(II) ions in the active site of SPU (PDB code 4CEU). (B) Superimposition of the open (white and light blue ribbons, PDB code 4CEU) and closed (grey and blue ribbons, PDB code 3UBP) conformations of the flexible flap in SPU, highlighting the side chain of αLys220* and αCys322.

water molecule, leading to a pseudo-square pyramidal penta-coordinated Ni1 and a pseudo-octahedral hexa-coordinated Ni2. A third water molecule completes, together with the bridging hydroxide, a cluster of solvent-derived O atoms that defines a tetrahedral cavity that is assumed to stabilize the transition state of the catalysed reaction.[4] The active site is found at the end of a *ca.* 15 Å-deep pocket characterized by the presence of a helix-turn-helix flap located at its entrance (Fig. 1B and 1-ESI†). This flap was observed to adopt different conformations, and has been suggested to modulate both the access of substrate to the enzyme active site and the dimensions of the catalytic cavity.[4a,5]

The design and synthesis of novel and potent urease inhibitors necessary to modulate its catalytic activity to counterbalance its negative effects require the knowledge of the structural details of the inhibition mechanism. Several urease structures reveal how the Ni(II) ions in the active site could be targeted by competitive inhibitors.[1b,6] On the other hand, the most efficient urease inhibitors known so far contain the

aLaboratory of bioinorganic Chemistry, Department of Pharmacy and Biotechnology, University of Bologna, Viale G. Fanin 40, I-40127 Bologna, Italy.
E-mail: stefano.ciurli@unibo.it
bEuropean Molecular Biology Laboratory, DESY, Notkestraße 85, 22607 Hamburg, Germany

†Electronic supplementary information (ESI) available: Materials and methods, structural parameters and unbiased electron density maps of BQ inhibited SPU, details of the SPU surface, and the results of the quantum mechanical calculations. See DOI: 10.1039/c6dt00652c
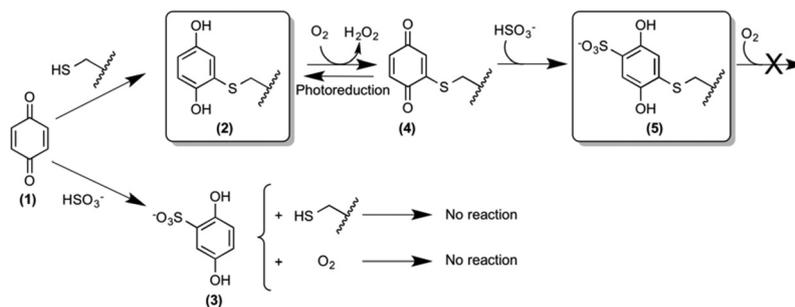
Scheme 1 Reactivity of 1,4-benzoquinone with the thiol group of solvent-exposed SPU cysteine residues in the absence and presence of sulphite. The structures of compounds **2** and **5** have been established by X-ray diffraction crystallography. The wiggled line represents the protein surface on which the reactive cysteines are located.

1,4-benzoquinone moiety (BQ, **1**, Scheme 1).[7] Quinones are known to react with thiols, leading to the formation of the corresponding thiol-substituted benzene-1,4-diol.[8] The sequence of the urease mobile flap contains the highly conserved residue αCys322 (residue numbering as in SPU), which is crucial for the activity of urease,[9] and whose position with respect to the active site is also modulated by the flap movement (Fig. 1B). Covalent modifications of this residue by the formation of a disulphide bond in the presence of 2-mercaptoethanol, demonstrated by crystallography,[10] lead to enzyme inhibition.[11] These observations support the hypothesis that the conserved cysteine in the flap is the target for quinone-based inhibitors, as shown in Scheme 1, by the formation of **2**.[7d] Here, we describe a kinetic characterization of the inhibition of SPU by 1,4-benzoquinone, and report, for the first time, the high-resolution crystal structure of the urease–BQ covalent complex, shedding light on the long-awaited molecular details of this interaction.

Kinetic measurements (see the ESI†) were performed at room temperature by using a spectrophotometric assay that monitors the decrease of urease activity as a function of time in the presence of increasing amounts of BQ (Fig. 2). The data were optimally fitted to a single exponential decay, and the resulting time constant k_{app} (s^{-1}) showed a linear dependence on the concentration of BQ (inset of Fig. 2). This behaviour is consistent with an irreversible process that has two reactants (SPU and BQ) and is first-order with respect to both of them, leading to an inactive form of SPU, hypothetically represented by compound **2** in Scheme 1:

$$BQ + SPU_{active} \xrightarrow{k} SPU_{inactive}$$
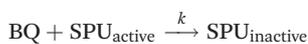
$$v = k\,[BQ]\,[SPU_{active}] = k_{app}\,[SPU_{active}];\ k_{app} = k\,[BQ]$$

The value for the second-order kinetic constant resulting from the linear fit is $k = 1.24 \pm 0.06 \times 10^3$ M^{-1} s^{-1}. The intercept of the linear fit in the inset of Fig. 2 crosses the abscissa at a value of 18.4 ± 1.5 μM; this value indicates the minimum concentration of BQ above which an inhibition effect can be detected. This observation was interpreted by considering the
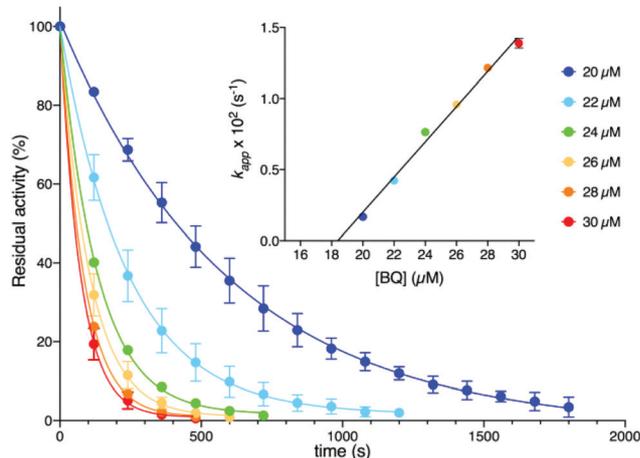
Fig. 2 SPU residual activity *vs.* time plot at different BQ concentrations. The inset shows a plot of k_{app} as a function of the BQ concentration. In both panels, the lines represent the result of an exponential or linear fit of the data.

presence of *ca.* 20 μM sodium sulphite in the enzymatic assay solution, added as a stabilizer of urease[6a] but also known to react stoichiometrically with BQ to give 2,5-dihydroxy-benzene-sulphonate (**3**, Scheme 1).[12] The evident lack of reactivity between **3** and SPU, compared to the reactivity of **1**, could be ascribed to the increase of the reduction potential caused by the electron withdrawing effects of the sulphonate moiety, which prevents the oxidation of **3** by dissolved molecular oxygen to the quinone level necessary for the reaction with the cysteine thiol (Scheme 1).[12,13]

In order to elucidate the structural basis of urease inactivation by BQ, SPU was incubated with **1** for two hours in the absence of sulphite in solution before crystallization. The resulting 2.07 Å crystal structure of SPU inhibited by BQ (see Table 1 and the ESI† for full details of crystallization, X-ray data collection, elaboration and analysis) confirmed the well-established architecture of the enzyme, as well as the rigidity of the protein scaffold, with a root mean square deviation (RMSD) for the Cα residue atoms of only 0.12 Å with respect to the structure of native urease (PDB code 4CEU).

View Article Online

Dalton Transactions

Communication

Published on 02 March 2016. Downloaded by UNIVERSITA DEGLI STUDI BOLOGNA on 20/03/2016 17:40:44.

Table 1 X-ray diffraction data collection and refinement statistics for BQ-bound SPU in the absence (PDB code 5FSE) and in the presence (PDB code 5FSD) of sulphite

Data collection	5FSE (BQ) @ 2.07 Å	5FSD (BQ + HSO₃⁻) @ 1.75 Å
Wavelength (Å)	0.954	1.214
Crystal-to-detector distance (mm)	378.1	256.9
Oscillation angle (°)	0.10	0.10
Number of images	400	1200
Space group	$P6_322$	$P6_322$
Unit cell (a, b, c, Å)	131.83, 131.83, 188.58	131.31, 131.31, 188.92
Resolution range[a] (Å)	188.58–2.07 (2.13–2.07)	97.43–1.75 (1.78–1.75)
Total number of reflections[a]	257 019 (18 993)	1 243 307 (51 825)
Unique reflections[a]	58 491 (4412)	96 874 (4696)
Redundancy[a]	4.4 (4.3)	12.8 (11.0)
Completeness[a] (%)	98.8 (97.6)	100.0 (100.0)
R_{sym} [a,b] (%)	13.5 (90.8)	10.1 (145.4)
R_{pim} [a,c] (%)	8.8 (55.1)	4.2 (66.5)
Mean I half-set correlation CC(1/2)	0.620	0.723
Mean $I/\sigma(I)$ [a]	9.0 (1.6)	20.0 (1.6)
Refinement statistics		
Number of urease subunits in the asymmetric unit	3	3
R_{factor} [d] (%)	14.4	13.4
R_{free} [d] (%)	19.1	16.9
Cruickshank's DPI for coordinate error[e] based on R_{factor} (Å)	0.15	0.08
Wilson plot B-factor (Å²)	22.6	21.5
Average all atom B-factor[f] (Å²)	30.77	29.16
Cys322 B-factor[f] (Å²) (Cβ, Sγ)	63.2, 76.5	52.4, 61.7
Cys555 B-factor[f] (Å²) (Cβ, Sγ)	37.9, 42.2	39.4, 42.1
Average ligand – Cys322 B-factor[f] (Å²)	66.5	55.6
Average ligand – Cys555 B-factor[f] (Å²)	41.1	43.1
RMS (bonds)[d]	0.016	0.019
RMS (angles)[d]	1.803	1.973
Total number of atoms[c]	6800	7012
Total number of water molecules	465	642
Solvent content (%)	55.50	55.23
Matthews coefficient (Å³/Da)	2.76	2.75
Ramachandran plot[g]		
Most favoured region (%)	90.1	91.6
Additionally allowed region (%)	9.0	7.5
Generously allowed region (%)	0.6	0.8
Disallowed region (%)	0.3	0.2

[a] Highest resolution bin in parentheses. [b] $R_{sym} = \sum_{hkl} \sum_j |I_j - <I>| / \sum_{hkl} \sum_j I_j$, where I is the intensity of a reflection, and $<I>$ is the mean intensity of all symmetry related reflections j. [c] $R_{p.i.m.} = \sum_{hkl} \{[1/(N-1)] \cdot \sum_j |I_j - <I>|\}^{1/2} / \sum_{hkl} \sum_j I_j$, where I is the intensity of a reflection, and $<I>$ is the mean intensity of all symmetry related reflections j, and N is the multiplicity.[15] [d] Taken from REFMAC;[16] R_{free} is calculated using 5% of the total reflections that were randomly selected and excluded from refinement. [e] DPI = $R_{factor} \cdot D_{max} \cdot compl^{-1/3} \cdot [N_{atoms}/(N_{refl} - Nparams)]^{1/2}$, where N_{atoms} is the number of the atoms included in the refinement, N_{refl} is the number of the reflections included in the refinement, D_{max} is the maximum resolution of reflections included in the refinement, $compl$ is the completeness of the observed data, and for isotropic refinement, $N_{params} \approx 4N_{atoms}$.[17] [f] Taken from BAVERAGE.[18] [g] Taken from PROCHECK.[19]

No significant changes were observed for the Ni(II) coordination environment (Table 1-ESI†). The mobile flap covering the active site cavity is found in the open conformation, as in all SPU structures determined so far with the notable exception of the complex with diamidophosphate (DAP), an analogue of the intermediate or the transition state of the reaction.[4a] The unbiased omit electron density map calculated with Fourier coefficients F_o–F_c and phases refined using the model of the citrate-inhibited SPU (PDB code 4AC7, 1.50 Å resolution[14]), shows an additional electron density in the vicinity of the flap and proximal to the αCys322 residue (Fig. 2A-ESI†). This electron density could be successfully modelled with full occupancy using the aromatic dihydroxylated moiety of the inhibitor, as shown by the $2F_o$–F_c map (Fig. 3A). The final refined structure features a covalent bond between αCys322 Sγ and an aromatic C atom of the inhibitor, with a C–S distance of 1.64 Å (Fig. 3B). The same covalent modification is also observed for αCys555, a residue located on the protein surface far from the active site, with a similar C–S distance of 1.67 Å (Fig. 3-ESI†). The remaining αCys520, protected from the solvent, is not affected by this modification.
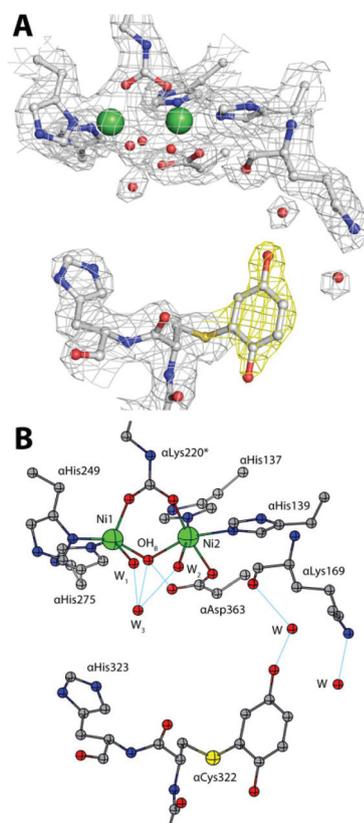
Fig. 3 An atomic model of the active site of BQ-inactivated SPU (PDB code 5FSE) in the absence of sulphite. In panel (**A**), the nickel-coordination environment is shown superimposed on the final $2F_o$–F_c electron density map contoured at 1σ; the map of the inhibitor is shown in yellow. In panel (**B**), the corresponding structural model is shown. H-bonds are shown as thin blue lines.

This journal is © The Royal Society of Chemistry 2016

Dalton Trans.

This result finally provides full structural support to the mechanism by which the solvent-exposed side chain thiols of cysteine residues in ureases are the reactive groups involved in the formation of covalent adducts with BQ.[7d]

In the case of αCys322, the O atom in the *meta* position with respect to the thiol substituent points towards the entrance of the active site channel (Fig. 3B), forming a hydrogen bond with a water molecule (at 2.85 Å) that is also stabilized by H-bonds with αLeu365 O (at 2.66 Å) and αLys169 O (at 2.76 Å).

The 1.75 Å resolution structure of SPU obtained from the crystals generated by adding 50 mM sodium sulphite to the crystallization drop obtained as before discloses further details of the reactivity of the enzyme inactivated by BQ. In this case, while the overall protein structure is maintained (as revealed by a value of 0.10 Å for the RMSD with respect to the structure obtained in the absence of sulphite, and no changes around the Ni(II) ions, see Table 1-ESI†), an additional electron density was identified in the unbiased omit F_o–F_c map calculated using the refined model of the 2.07 Å structure described above, in the proximity of the inhibitor aromatic ring bound to αCys322 (Fig. 2B-ESI†). This additional electron density was successfully modelled as a sulphonate group covalently bound to the aromatic ring in the *para* position with respect to the cysteine thiol, with 70% occupancy (Fig. 4A). One of the O atoms of the sulphonate group forms a H-bond with αLys169 Nζ (at 2.85 Å, Fig. 4B). The same covalent modification is also observed for αCys555, albeit with a lower occupancy (Fig. 4-ESI†). A comparative analysis of the solvent-excluded surfaces of the enzyme active site channel, in the absence and presence of the covalent modifications of αCys322 with and without sulphite in the crystallization milieu (Fig. 5-ESI†) shows that the channel is not physically closed in the presence of the bound inhibitor, independently of the sulphonate modification, suggesting that the inactivation of the enzyme is not caused by the blockage of substrate transport through the path leading into the active site, but it is rather due to the prevention of the flap to close, as necessary for the catalysis to occur.[5] A site-directed mutation of the conserved cysteine with tyrosine in *K. aerogenes* urease, supposedly having a similar effect on the active site entrance, causes the abolition of catalytic activity.[9]

The oxidation state (BQ or 1,4-hydroquinone, HQ) of the inhibitor bound to αCys322 and αCys555 was also a matter of investigation. In order to discern between these two possibilities, both moieties were used as initial models for crystallographic refinement, yielding a negligible difference in data statistics. Quantum-mechanical calculations were thus carried out on the two possible redox moieties (see the ESI†): a comparison of the key C–O distances for the covalently modified urease cysteines obtained by crystallography with those derived from these calculations (see Tables 3, 4-ESI and Fig. 6, 7-ESI†), did not allow us to unequivocally discern between the oxidized or reduced form in the crystal. However, a comparison of the C1–C2–Sγ–Cβ dihedral angle (see Scheme 1-ESI†) suggests the presence of the reduced form, characterized by a value close to 90° both in the crystal struc-
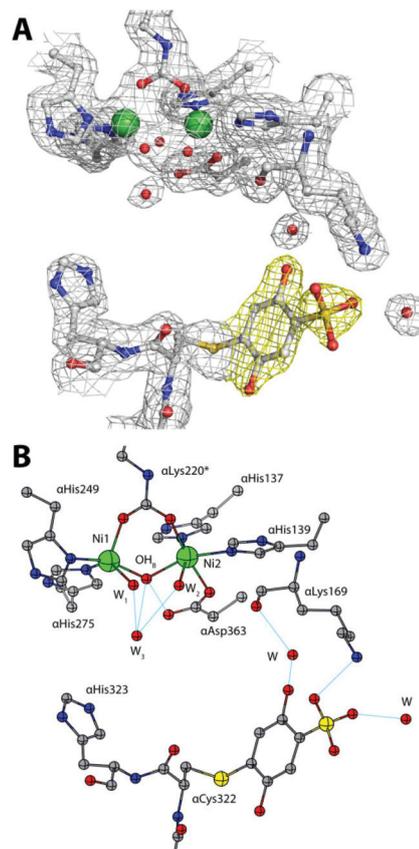
Fig. 4 An atomic model of the active site of BQ-inactivated SPU (PDB code 5FSD) in the presence of sulphite. (**A**) The nickel-coordination environment superimposed on the final $2F_o$–F_c electron density 1σ map; (**B**) the corresponding structural model. The figure settings are the same as in Fig. 3.

ture and in the theoretical model (see Tables 5, 6-ESI and Fig. 6, 7-ESI†). This particular conformation is known to be due to the formation of an intra-molecular hydrogen-bond between a S atom p-orbital and the O–H group of the HQ.[20] The final refinement for both structures was thus carried out using the reduced form of BQ. The reaction of the BQ-inactivated enzyme with sulphite in solution, leading to the crystallographically characterized compound **5** (Scheme 1), requires however the presence of the oxidized form **4**, which must be obtained from the initial product **2** by reaction with dissolved molecular oxygen (Scheme 1). The presence of the reduced form **2** in the solid state could thus be due to photo-reduction. In the case of the sulphonate-substituted compound **5**, on the other hand, the presence of the reduced form is consistent with its known stabilization by the electron-withdrawing properties of the sulphonate group, as discussed for compound **3**.

In this paper we have described and discussed the reactivity of αCys322, a key residue for the catalytic activity of urease, towards 1,4-benzoquinone, and established the molecular structural details of the resulting irreversibly inactivated enzyme. This information is essential for the development of

structure-based design and structure–activity relationship studies of novel urease inhibitors based on this moiety. Moreover, considering the conservation of the cysteine on the mobile flap of ureases from other sources, such as the pathogenic bacterium *Helicobacter pylori*, the results of our study have wider general applications. Finally, the large class of quinonoid compounds, containing **1** as their basic core, is widely distributed in nature and play pivotal roles as electron transfer agents in primary metabolic processes such as photosynthesis, respiration and oxidative phosphorylation.[21] A large number of chemical derivatives of differently substituted quinones are also involved in pharmacological applications for their activity as antibiotics, antitumor, antimalarial, antineoplastic, anticoagulant and herbicidal agents. Therefore, the structural information obtained in our study could have far reaching consequences in the comprehension of the reactivity of this class of compounds with thiol groups in proteins, enzymes and enzyme cofactors.[22]

Notes and references

1 (a) B. Zambelli, F. Musiani, S. Benini and S. Ciurli, *Acc. Chem. Res.*, 2011, **44**, 520–530; (b) M. J. Maroney and S. Ciurli, *Chem. Rev.*, 2014, **114**, 4206–4228.
2 J. C. Rutherford, *PLoS Pathog.*, 2014, **10**, e1004062.
3 S. Kiss and M. Simihaian, *Improving efficiency of urea fertilizers by inhibition of soil urease activity*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.
4 (a) S. Benini, W. R. Rypniewski, K. S. Wilson, S. Miletti, S. Ciurli and S. Mangani, *Structure*, 1999, **7**, 205–216; (b) S. Ciurli, S. Benini, W. R. Rypniewski, K. S. Wilson, S. Miletti and S. Mangani, *Coord. Chem. Rev.*, 1999, **190–192**, 331–355.
5 F. Musiani, E. Arnofi, R. Casadio and S. Ciurli, *J. Biol. Inorg. Chem.*, 2001, **6**, 300–314.
6 (a) L. Mazzei, M. Cianci, S. Benini, L. Bertini, F. Musiani and S. Ciurli, *J. Inorg. Biochem.*, 2016, **154**, 42–49; (b) S. Benini, M. Cianci, L. Mazzei and S. Ciurli, *J. Biol. Inorg. Chem.*, 2014, **19**, 1243–1261.
7 (a) J. M. Bremner and L. A. Douglas, *Soil Biol. Biochem.*, 1971, **3**, 297–307; (b) R. L. Mulvaney and J. M. Bremner, *Soil Biol. Biochem.*, 1978, **10**, 297–302; (c) W. Zaborska, M. Kot and K. Superata, *J. Enzyme Inhib. Med. Chem.*, 2002, **17**, 247–253; (d) W. Zaborska, B. Krajewska, M. Kot and W. Karcz, *Bioorg. Chem.*, 2007, **35**, 233–242; (e) M. Kot and W. Zaborska, *J. Enzyme Inhib. Med. Chem.*, 2006, **21**, 537–542; (f) B. Krajewska, *J. Mol. Catal. B: Enzym.*, 2009, **59**, 9–21.
8 J. M. Snell and A. Weissberger, *J. Am. Chem. Soc.*, 1939, **61**, 450–453.
9 P. R. Martin and R. P. Hausinger, *J. Biol. Chem.*, 1992, **267**, 20024–20027.
10 S. Benini, W. R. Rypniewski, K. S. Wilson, S. Ciurli and S. Mangani, *J. Biol. Inorg. Chem.*, 1998, **3**, 268–273.
11 M. J. Todd and R. P. Hausinger, *J. Biol. Chem.*, 1989, **264**, 15835–15842.
12 T. H. James, J. M. Snell and A. Weissberger, *J. Am. Chem. Soc.*, 1938, **60**, 2084–2093.
13 (a) T. H. James and A. Weissberger, *J. Am. Chem. Soc.*, 1938, **60**, 98–104; (b) T. H. James and A. Weissberger, *J. Am. Chem. Soc.*, 1939, **61**, 442–450.
14 S. Benini, P. Kosikowska, M. Cianci, L. Mazzei, A. G. Vara, L. Berlicki and S. Ciurli, *J. Biol. Inorg. Chem.*, 2013, **18**, 391–399.
15 M. S. Weiss, *J. Appl. Crystallogr.*, 2001, **34**, 130–135.
16 (a) G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta Crystallogr.*, 1997, **53**, 240–255; (b) G. N. Murshudov, A. A. Vagin, A. Lebedev, K. S. Wilson and E. J. Dodson, *Acta Crystallogr.*, 1999, **55**, 247–255.
17 D. W. J. Cruickshank, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1999, **55**, 583–601.
18 M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2011, **67**, 235–242.
19 R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *J. Appl. Crystallogr.*, 1993, **26**, 283–291.
20 P. Politzer, J. S. Murray and P. Lane, *Int. J. Quantum Chem.*, 2007, **107**, 3046–3052.
21 I. Abraham, R. Joshi, P. Pardasani and R. T. Pardasani, *J. Braz. Chem. Soc.*, 2011, **22**, 385–421.
22 W.-W. Li, J. Heinze and W. Haehnel, *J. Am. Chem. Soc.*, 2005, **127**, 6140–6141.

ANNEX 5

Reprinted from:

Journal: Journal of Inorganic Biochemistry

Title: Inactivation of urease by catechol: Kinetics and structure

Volume: 166

Year: 2017

Pages: 182 - 189

Authors: L. Mazzei, M. Cianci, F. Musiani, G. Lente, M. Palombo and S. Ciurli

Inactivation of urease by catechol: Kinetics and structure

Luca Mazzei [a], Michele Cianci [b,1], Francesco Musiani [a], Gábor Lente [c], Marta Palombo [a], Stefano Ciurli [a,*]

[a] Laboratory of Bioinorganic Chemistry, Department of Pharmacy and Biotechnology, University of Bologna, Italy
[b] European Molecular Biology Laboratory, DESY, Hamburg, Germany
[c] Department of Inorganic and Analytical Chemistry, University of Debrecen, Hungary

A B S T R A C T

Urease is a Ni(II)-containing enzyme that catalyzes the hydrolysis of urea to yield ammonia and carbamate at a rate 10^{15} times higher than the uncatalyzed reaction. Urease is a virulence factor of several human pathogens, in addition to decreasing the efficiency of soil organic nitrogen fertilization. Therefore, efficient urease inhibitors are actively sought. In this study, we describe a molecular characterization of the interaction between urease from *Sporosarcina pasteurii* (SPU) and *Canavalia ensiformis* (jack bean, JBU) with catechol, a model polyphenol. In particular, catechol irreversibly inactivates both SPU and JBU with a complex radical-based autocatalytic multistep mechanism. The crystal structure of the SPU-catechol complex, determined at 1.50 Å resolution, reveals the structural details of the enzyme inhibition.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Urease is a non-redox metallo-enzyme involved in the global nitrogen cycle in plants, algae, fungi, and several microorganisms [1,2]. Its fully conserved bimetallic Ni(II)-containing reaction site (Fig. 1A) catalyzes the rapid hydrolytic decomposition of urea to give ammonia and carbamate [1,2]. The uncatalyzed hydrolysis of these reaction products causes an overall pH increase that has negative implications for human health [3] as well as for the environmental ecosphere [4]. In particular, urease plays an essential role in the virulence of the human pathogen *Helicobacter pylori*. This bacterium is able to use the urease activity to create a local pH environment compatible with its survival in the acidic conditions of the gastric mucosa [5]. It has been estimated that approximately two-thirds of the world's population harbor *H. pylori* [6]. This microorganism has been classified, uniquely among bacteria, as a class-I carcinogen in humans [7]. As a consequence of these facts, urease is the ideal target for the treatment of diseases caused by a large variety of ureolytic human pathogens in addition to *H. pylori*, such as *Cryptococcus neoformans*, *Yersinia enterocolitica*, *Proteus mirabilis*, and *Mycobacterium tuberculosis* [8–11]. The urease inhibitors characterized so far [12] can be roughly divided in two main classes, based on the mode of interaction with the enzyme. The first class contains molecules (phosphorodiamidate, phosphate, thiols, sulfite, fluoride as well as hydroxamic, citric and boric acids) directly binding to the Ni(II) ions in the active site [1,2,13,14]. The second class of inhibitors is composed of molecules such as 2-mercapto-ethanol (BME) [15] and 1,4-benzoquinone (PBQ) [16] that have been proven to bind to the Sγ atom of the conserved αCys322 [*Sporosarcina pasteurii* urease (SPU) numbering]. This residue is located onto a mobile helix-turn-helix motif (flap) that is essential for catalysis by modulating the transit of substrate and products through the active site cavity (Fig. 1B). BME can actually be ascribed to both classes, because it is able to further use its thiol moiety to coordinatively bridge the Ni(II) ions in the active site [15]. On the other hand, PBQ is a representative of the most efficient class of urease inhibitors characterized so far [17]. The structure of SPU inactivated by PBQ shows that the thiol group of αCys322 forms a covalent adduct with PBQ, blocking enzyme activity by rendering the flap unable to move [16]. A similar inactivation mode could potentially represent the basis for the inhibitory activity of polyphenols towards urease. These molecules comprise a wide class of natural compounds found in fruits, vegetables, cereals, tea, coffee and wine, and are generally recognized as beneficial to human health for their antioxidant properties [18]. This class of urease inhibitors has also been recently described as efficient against *H. pylori*, but their mode of action has not been understood so far [19–22]. Here we present the kinetic and structural characterization of SPU and *Canavalia ensiformis* (jack bean, JBU) inactivation by catechol (benzene-1,2-diol, CAT), the simplest molecule featuring a polyphenol scaffold that was identified as the most promising inhibitor of soil urease [17].

* Corresponding author at: Laboratory of Bioinorganic Chemistry, Department of Pharmacy and Biotechnology, University of Bologna, Viale G. Fanin 40, I-40127 Bologna, Italy.
E-mail address: stefano.ciurli@unibo.it (S. Ciurli).
1 Present address: Department of Agricultural, Food and Environmental Sciences, Università Politecnica delle Marche, Ancona, Italy.
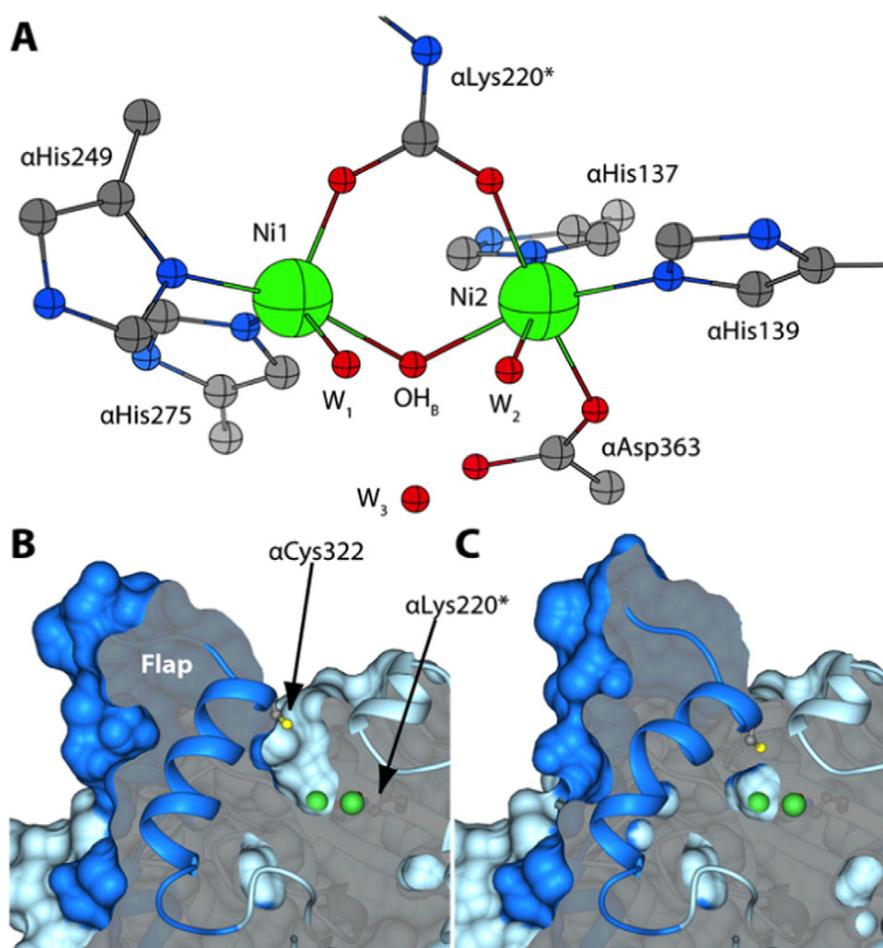
Fig. 1. (A) Coordination environment of the Ni(II) ions in the active site of SPU (PDB code 4CEU [13]). Ribbons and longitudinal section of the open (B, PDB code 4CEU) and closed (C, PDB code 3UBP [42]) conformations of the flexible flap in SPU.

2. Materials and methods

2.1. Enzymes

Sporosarcina pasteurii urease was expressed and purified from the native source, following a previously described protocol [14–16]. The protein was quantified by measuring the activity and considering its specific activity of 2500 units mg^{-1} and $M_r = 250$ kDa. Jack bean urease type C-3, powder (≥600,000 units/g) was purchased from Sigma-Aldrich and was quantified following manufacturer's information.

2.2. Kinetic measurements

Pre-incubation experiments were carried out at room temperature by using a spectrophotometric assay in which cresol red is exploited as a probe to follow the overtime increase of pH due to urease activity, following a protocol previously described [16]. In the case of SPU – catechol experiments, a solution of 50 nM SPU dissolved in 2 mM HEPES buffer, pH 7.5, 1 mM Na_2SO_3, was diluted down to 1 nM final concentration in the cresol red reaction solution, containing 30 mg L^{-1} cresol red in 2 mM HEPES buffer, pH 7.5, 2 mM EDTA. By doing this, Na_2SO_3 was diluted 50-fold as well, down to 20 μM final concentration. In the case of JBU – catechol experiments, a solution of 250 nM JBU dissolved in 2 mM HEPES buffer, pH 7.5, without Na_2SO_3, was directly diluted to 2.5 nM in the same cresol red solution. Then, different concentrations of catechol solutions (4–16 mM and 5–40 μM for the experiments on SPU and JBU, respectively) dissolved in 2 mM HEPES buffer, pH 7.5

(containing 20 μM Na_2SO_3 in the case of experiments on SPU) were added, taking the time when the enzyme solution and the ligand are mixed as zero time of incubation. After appropriate periods of time, aliquots were withdrawn from the incubation solution, 100 mM urea was added, and the change in absorbance over time was monitored (λ = 573 nm). The activity was calculated by a linear fitting of the straight portion in the absorbance vs. time curve and normalized to the activity measured at time zero incubation.

2.3. Crystallization, data collection and structural determination

A solution of *Sporosarcina pasteurii* urease (SPU, 1 mg mL^{-1}) in 20 mM HEPES buffer, pH 7.5, was incubated in the presence of 1 mM catechol (dissolved in the same buffer) for 2 h and then concentrated to 11 mg mL^{-1}. Subsequently, 2 μL of this solution was diluted with 2 μL of the precipitant solution (1.6–2.0 M ammonium sulfate in a 50 mM sodium citrate buffer pH 6.3). Crystallization was performed at 293 K using the hanging-drop method, equilibrating the drop against 0.5 mL of the precipitant solution using Quiagen EasyXtal 15-Well plates. Rice-shaped protein crystals appeared in 1–2 weeks and grew to a size of 0.1 × 0.1 × 0.3 mm^3. Crystals were scooped up using cryoloops, transferred to a cryoprotectant solution containing 20% ethylene glycol, 2.4 M ammonium sulfate in a 50 mM sodium citrate buffer at pH 6.3 and then flash-cooled and stored in liquid nitrogen. Diffraction data were collected at 100 K using synchrotron radiation at the EMBL P13 beamline of the Petra III storage ring, c/o DESY, Hamburg (Germany). Diffraction images were collected by performing helical scans

along the crystal to achieve higher resolution by minimizing radiation damage. Data were processed using XDS [23] and AIMLESS [24]. The crystals were isomorphous with those of native urease and other complexes of the same enzyme. The model of SPU with its highest resolution available so far, that is in complex with citrate (PDB code 4AC7, 1.50 Å resolution [25]) and devoid of water molecules and ligands, was used as a starting model for the rigid body refinement of the single αβγ protein trimer, carried out using Refmac [26]. Model building and water or ligand addition/inspection were conducted using Coot [27]. The structure was refined using isotropic atomic displacement parameters (ADPs), including the hydrogen atoms in the riding positions, and then refined using anisotropic ADPs. The model was then validated using the PDB_REDO web server [28]. The final R and R_{free} were 12.12 and 15.42, respectively. The diffraction data and final refinement statistics are given in Table S2. The structure was deposited in the Protein Data Bank under the accession code 5G4H. Figures were generated using UCSF Chimera [29], PyMol [30], and CrystalMaker (http://www.crystalmaker.com).

2.4. Quantum mechanical calculations

Density functional theory (DFT) computations were carried out using the program ORCA 3.0.3 [31] and the Becke three-parameter hybrid functional combined with Lee-Yang-Parr correlation functional (B3LYP/G) [32,33] as defined in the Gaussian software [34]. The formation of the bond between the C3 atom of *ortho*-benzoquinone (cyclohexa-3,5-diene-1,2-dione, OBQ) or *ortho*-semiquinone [(2-hydroxyphenyl)-oxidanyl, OSQ] and αCys322(Sγ) was investigated using relaxed scan computations [31], which involve constrained optimizations for different values of a selected reaction coordinate. The chosen coordinate, in this case the bond distance, is fixed to a certain value while the remaining coordinates of the molecule are optimized. After completion of one optimization cycle, the value of the reaction coordinate is changed and another optimization cycle is carried out. This procedure was used to trace a path (theoretically close to a minimum-energy path) from the separated OBQ (or OSQ) and a molecule of methanethiol, used to model αCys322 thiol, to the product. For relaxed scan calculations all atoms were described by the Pople-style 6-311G(d, p) [35] basis set. The Dunning correlation-consistent polarized double zeta basis set (cc-pVDZ) [36] was used to calculate the hydrogen bond network around the catechol moiety bound to the mobile flap of SPU, while the Dunning correlation-consistent polarized triple zeta basis set, with the inclusion of diffuse functions (aug-cc-pVTZ) [36], was used for computations on small molecules. Frequency computations were executed on the molecules (reported in Fig. S3 and Table S5) as well as on the critical points reported in Fig. 4, in order to determine their nature.

3. Results and discussion

The inhibition of *Sporosarcina pasteurii* urease (purified according to previously published protocols [14–16]) by catechol was studied by pre-incubating the enzyme with different concentrations of inhibitor for increasing periods of time in the absence of substrate, and the residual activity was monitored using a spectrophotometric assay [16]. The data in Fig. 2A show the presence of an initial lag phase, whose extent decreases as the concentration of catechol increases, followed by an acceleration that eventually yields the complete inhibition of the enzyme. This behavior is typical for irreversible enzyme inactivators, known to form stable covalent adducts between the inhibitor and essential functional groups of the enzyme [37].

The observed kinetic response suggests the presence of a process that converts catechol, inactive at the start of the reaction, into another compound, which then acts as the actual inactivator of the enzyme. Catechol is known to undergo oxidation in neutral aqueous solutions upon exposure to air [38]. The need for dissolved molecular oxygen was
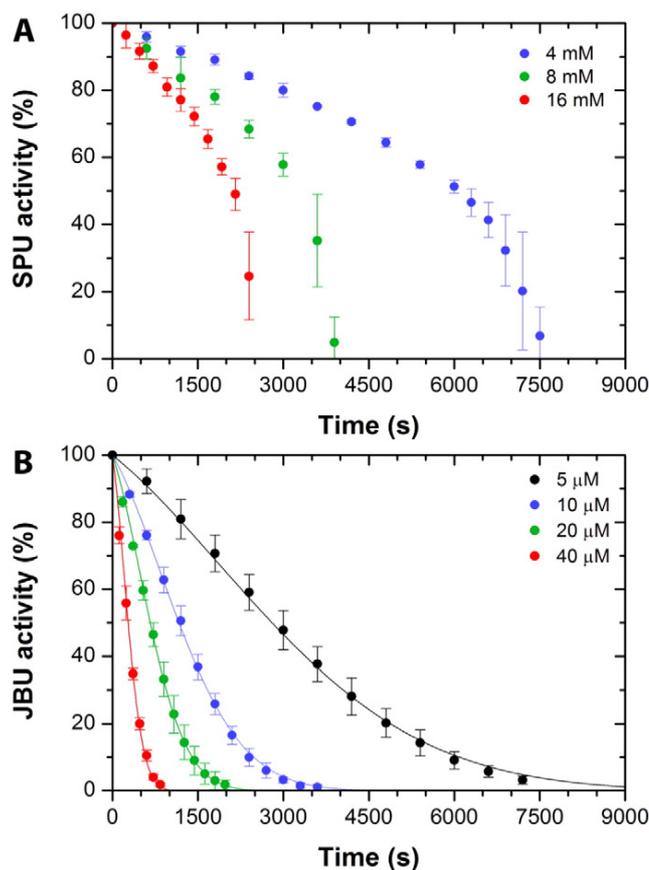
Fig. 2. Plots of SPU (A) and JBU (B) residual activity vs. time of incubation at different catechol concentrations. In panel B, the lines represent the result of the global fits of the data for JBU using Eq. (1) (see below).

confirmed by the essential lack of SPU inactivation when the kinetic experiments described above were performed in controlled oxygen-free atmosphere (Fig. S1 in the Supplementary material).

According to the known reduction potentials (see Table S1), the formation of the *ortho*-semiquinone radical or of *ortho*-benzoquinone are thermodynamically favorable processes, as is the disproportionation of the *ortho*-semiquinone to form *ortho*-benzoquinone and catechol. These reactions are expected to be much faster than the observed lag phase [39], which therefore cannot be ascribed to these processes. We thus investigated the possibility that sulfite, needed in the assay solution to stabilize SPU activity [14], could quench the urease inactivation by acting as a reducing agent of either OBQ or OSQ, or both, as suggested by the reduction potentials shown in Table S1. In addition, sulfite can also react with quinones, forming dihydroxybenzene-sulphonates shown to be inactive towards the cysteine thiols of SPU [16]. In order to explore this hypothesis, the same kinetic experiments were conducted using, in place of SPU, *Canavalia ensiformis* urease (jack bean urease, JBU), which does not require the presence of sulfite as a preservative. Fig. 2B shows that, in this latter case, the initial lag phase is almost completely abolished, supporting the role of sulfite in slowing down the inactivation of SPU. In addition, the concentration of catechol needed to exert similar effects on urease activity is three orders of magnitude smaller in the case of JBU as compared to that needed for SPU. The sigmoidal shape of the JBU inactivation data as a function of incubation time suggests the occurrence of a collectively autocatalytic process, namely a set of chemical reactions that yield, as reaction products, one or more catalysts that accelerate the inhibition of JBU [40]. The nature of this autocatalytic process for JBU inactivation must involve the production of derivatives of catechol generated by reaction with dioxygen,

namely OSQ and/or OBQ. We speculate here that a similar autocatalytic process also occurs for SPU, but only upon consumption of sulfite, when the fast phase of SPU inactivation is initiated by unrestrained formation of the oxidation products of catechol.

The derivation of a complete kinetic model able to interpret the data requires the identification of the reacting species and of the reaction product(s). Previous studies on the inactivation of JBU by catechol were interpreted with the covalent modification of αCys592 (corresponding to αCys322 in SPU) through a Michael-type nucleophilic addition by its thiol group on OBQ, even though a radical-based mechanism was not ruled out [41]. In the most general case, this residue could react with either OSQ or OBQ, depending on the nature of the reaction mechanism: a nucleophilic Michael-type addition on OBQ or a radical addition on OSQ. Furthermore, two different products could be a priori obtained, depending on the formation of a S—C covalent bond in ortho (C3 carbon) or in meta (C4 carbon) with respect to the hydroxyl groups on the aromatic ring. In order to understand the structural basis of urease inactivation by catechol, a crystallographic structural study of the products of inactivation of SPU by this inhibitor was carried out.

SPU was incubated with catechol for two hours before crystallization. The resulting crystals were amenable to X-ray diffraction analysis, yielding a 1.50 Å resolution structure (see Table S2 for the full details of X-ray data collection, processing and analysis). The structure confirms the well-known molecular architecture of SPU together with the rigidity of the protein scaffold (root mean square deviation for the Cα residue atoms = 0.08 Å with respect to the structure of native urease; PDB code 4CEU). The coordination environment of Ni(II) is almost identical to that observed in the case of native SPU (Table 1 and Table S3). The mobile flap covering the active site cavity is found in the open conformation, as also found in all SPU structures determined so far, with the distinguished exception of the complex with diamidophosphate (DAP), an analogue of the intermediate or transition state of the reaction [42].

The unbiased omit electron density map was calculated with Fourier coefficients F_o–F_c and phases derived from the refinement of SPU-catechol structure using the citrate-inhibited SPU (PDB code 4AC7, 1.50 Å resolution [25]) as a starting model after removal of the citrate moiety and other water molecules/ligands present in the 4AC7 model. This omit map showed an additional electron density in the vicinity of the flap and proximal to the αCys322 residue (Fig. S2), which was successfully modeled with full occupancy using the aromatic dihydroxylated moiety of the inhibitor (Fig. 3A). The final refined structure features a covalent bond between αCys322 Sγ and the aromatic C3 atom of the inhibitor, with a C—S distance of 1.65 Å (Fig. 3B and Table S4). The dihydroxylated aromatic moiety points towards the Ni(II) ions. The solvent-exposed αCys555, far from the active site, is not modified by catechol, as observed in the case of PBQ bound SPU [16], emphasizing the peculiarity of the environment around αCys322. This feature could be

Table 1
Selected distances around the Ni(II) ions in native (PDB code 4CEU) and catechol-bound SPU (PDB code 5G4H).

Ni - L distances (Å)	4CEU	5G4H
Ni(1) - αLys220* Oθ1	1.94	2.02
Ni(1) - O(B)[a]	2.08	1.90
Ni(1) - O(1)	2.24	2.21
Ni(1) - αHis249 Nδ	2.03	2.04
Ni(1) - αHis275 Nε	2.02	2.09
Ni(2) - αLys220* Oθ2	2.08	2.09
Ni(2) - O(B)	2.12	1.97
Ni(2) - O(2)[c]	2.07	2.17
Ni(2) - αHis137 Nε	2.11	2.10
Ni(2) - αHis139 Nε	2.08	2.10
Ni(2) - αAsp363 Oδ1	2.10	2.13
Ni(1) ⋯ Ni(2)	3.67	3.57
O(1) ⋯ O(2)	2.37	2.33
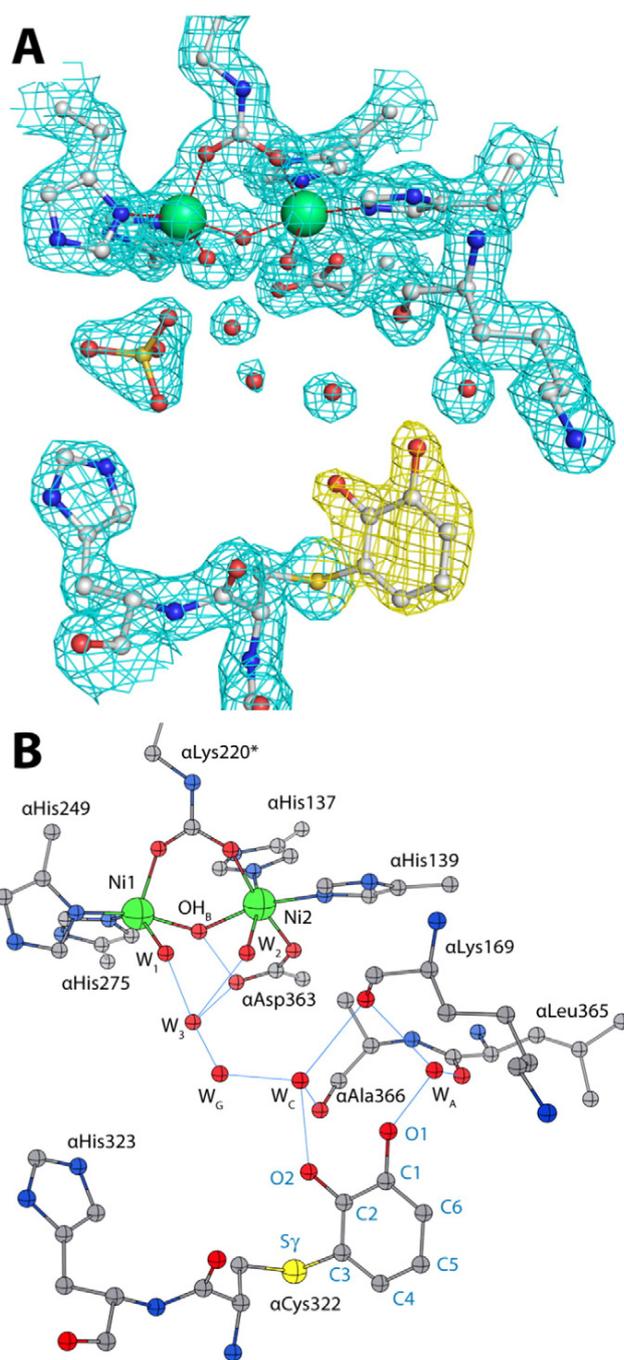
Fig. 3. Atomic model of the active site of catechol-inactivated SPU (PDB code 5G4H). In panel (A), the nickel-coordination environment is shown superimposed on the final $2F_o$–F_c electron density map contoured at 1.2σ; the map of the inhibitor is shown in yellow. In panel (B), the corresponding structural model is shown. H-bonds are shown as thin blue lines. Table S7 reports the corresponding numbers of the water molecules in the deposited PDB file.

related to the presence of the αHis323 residue in the adjacent position, which is conserved and known to be essential for urease activity. A third cysteine residue (αCys520) is protected from the solvent and is not affected by any modification.

The oxidation state of the group bound to αCys322 Sγ (reduced catechol or oxidized ortho-benzoquinone) was investigated using quantum-mechanical calculations carried out on these two possible redox moieties (see Supplementary material for details). The results suggest

the presence of the reduced form, and the final crystallographic refinement was thus carried out accordingly. In order to assess the position of the hydrogen atoms around the catechol moiety covalently bound to αCys322 Sγ, attention was focused on the network of H-bonds involving this modified residue and the water molecules in its proximity (see Supplementary material for details). In particular, W_A and W_B are bound to CAT O1 with O—O distances of 2.52 and 3.05 Å, respectively, while W_C is at H-bond distance from both CAT O1 and O2 (at 3.14 and 2.54 Å, respectively). W_A also forms two H-bonds with the backbone O atoms of αLys169 and αLeu365. W_C is at H-bond distance from both αAla366 O and W_G, and the latter is H-bonded to the sulfate ion located next to the Ni(II) ions. Quanto-mechanical optimization of the H-bonding network between the catechol moiety, the surrounding residues and water molecules suggest that W_A is actually the hydronium H_3O^+ ion; this induces the H atoms of the hydroxyl groups of catechol to lie on the same plane of the aromatic ring, forming an intra-molecular H-bond between a αCys322 Sγ p-orbital and the O—H group bound to the C2 atom of catechol [43].

The present crystal structure of SPU inactivated by catechol, with the αCys322 thiol bound to the C3 atom of the aromatic ring, rules out the formation of the *meta* adduct (cysteine thiol bound to the C4 atom of catechol). In order to distinguish between the nucleophilic attack on OBQ and the radical-based reaction mechanism involving OSQ, quanto-mechanical calculations were carried out (see Supplementary material for details). In both cases, the reaction involves the formation of a bond between OBQ/OSQ(C3) and αCys322(Sγ). The first mechanism taken into account was the nucleophilic attack performed by the methanethiol (MeSH) sulfur atom to the C3 atom of OBQ. The relaxed surface scan for this reaction coordinate leads to a gradual energy increase as the OBQ(C3)-MeSH(S) distance is shortened, without the formation of any critical point on the potential energy surface (Fig. S5). This result is somehow expected, considering that the C3 atom of OBQ has a slightly negative partial charge (Table S5), not compatible with a nucleophilic attack. We then took into account the direct interaction between OSQ(C3) and MeSH(S) (blue pathway in Fig. 4). The surface scan along the OSQ(C3)-MeSH(S) coordinate resulted in a barrier of ca.
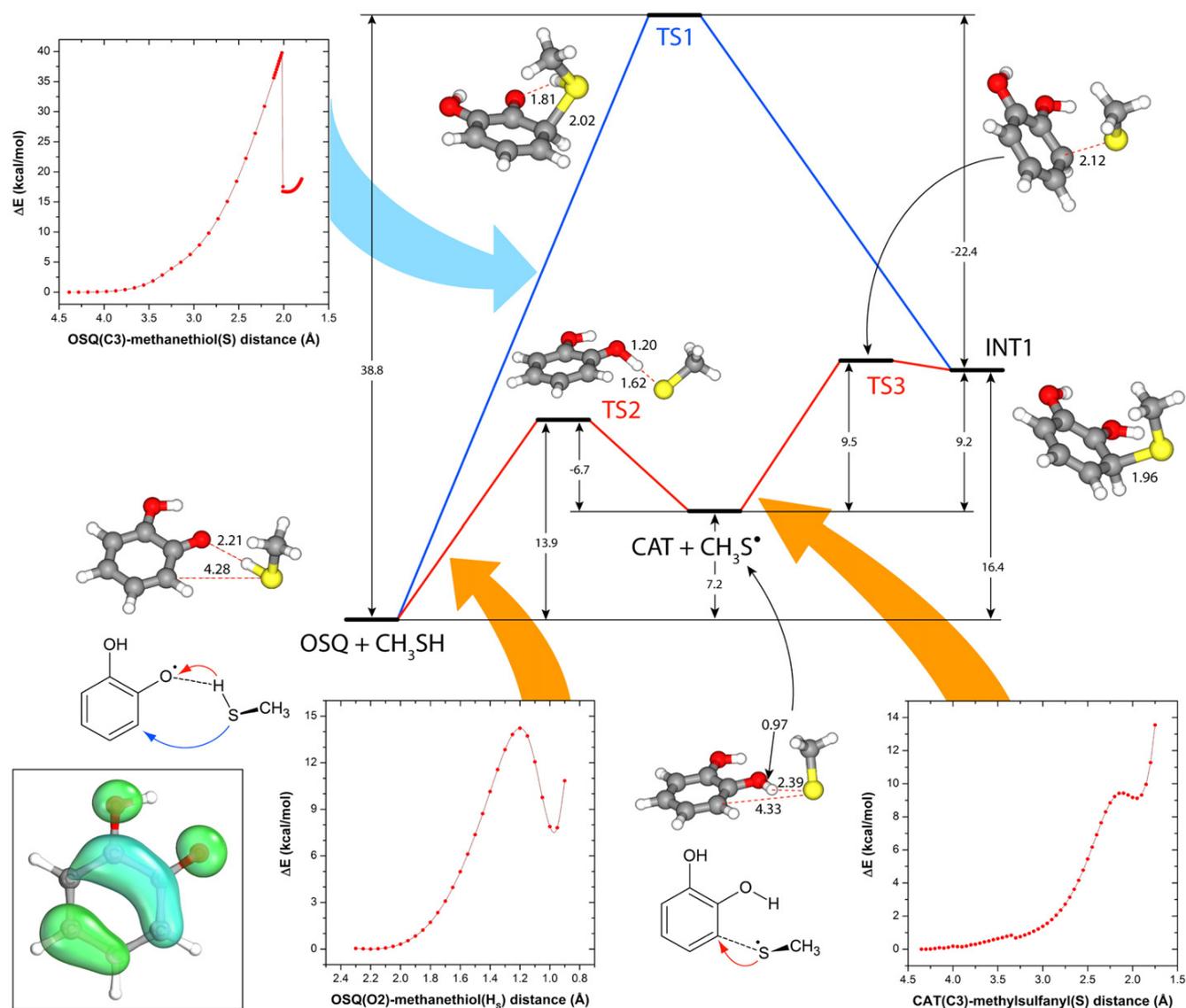
Fig. 4. Results of the relaxed surface scan computations and critical points identified along the reaction pathways between OSQ and MeSH. Atoms are colored accordingly to atom type. The inset reports the spin density calculated for the OSQ moiety.
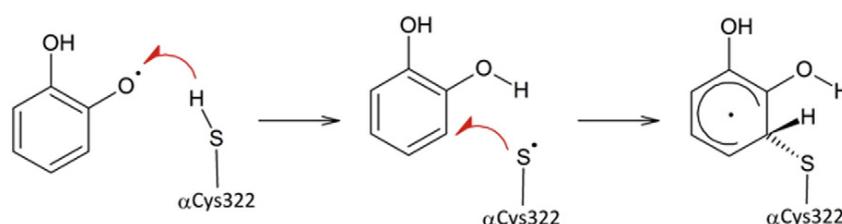
Scheme 1. Proposed reaction pathway for the SPU inactivation by OSQ leading to the formation of INT1, based on quanto-mechanical calculations.

39 kcal/mol and the identification of a transition state (TS1) characterized by C3 in a distorted tetrahedral geometry, together with a larger MeSH S—H bond distance as compared to that found in MeSH. Along this pathway, TS1 then evolves to a radical reaction intermediate (INT1) characterized again by C3 in a distorted tetrahedral geometry and full protonation of the two hydroxyl groups on the aromatic ring. The detailed analysis of the molecular orbitals of OSQ subsequently showed that the largest fraction of the spin density is localized on the O2 atom (Table S5), suggesting that the reaction could actually occur in two steps (Scheme 1), as previously suggested [44,45]: i) in the first step the thiol H (H_S) is taken up by the O2 of OSQ to form a molecule of catechol and a methylsulfanyl radical (MeS•), and ii) the S atom of MeS• then attacks the C3 of catechol to give the reaction products (red pathway in Fig. 4). The exploration of the OSQ(O2)-MeSH(H_S) reaction coordinate (Fig. 4) revealed a first energy maximum located at ca. 14 kcal/mol higher than the initial molecules, leading to the identification of the transition state TS2 characterized by the MeSH hydrogen already bound to OSQ O2 atom. TS2 then evolves to yield catechol and the MeS• radical. The exploration of the CAT(C3)-MeS•(S) reaction coordinate then leads to a second energy maximum characterized as the transition state (TS3) located at ca. 9.5 kcal/mol higher than the adduct formed by catechol and the MeS• radical. TS3 features the C3 atom of catechol in a distorted tetrahedral geometry. TS3 then evolves into the

same radical reaction intermediate INT1 observed in the reaction pathway initially explored, with very similar geometry and energy with respect to TS3. In both pathways, INT1 can be eventually stabilized by a third molecule able to extract a radical H atom and give the reaction product observed in the crystal structure of catechol-bound SPU. This third player could be either molecular oxygen or another molecule of OSQ (to give the fully reduced catechol-like moiety), or OBQ (to give OSQ). These calculations, coupled with the structural data, strongly support the involvement of the ortho-semiquinone radical in the reaction with the active form of urease to give the inactive enzyme.

The working hypothesis for urease inactivation by catechol thus implicates the set of reactions portrayed in Scheme 2. The OSQ radical can be generated by oxidation of catechol by dissolved O_2 through the formation of superoxide (reaction 1) and by comproportionation of catechol and ortho-benzoquinone (reaction 2). The active form of urease (JBU_{act} in the case of JBU) then reacts with OSQ to give the inactive enzyme (JBU_{inh}) according to the reaction mechanism based on quanto-mechanical calculations, as described above (reaction 3). The final reaction 4 takes into account a step that produces ortho-benzoquinone.[a]

This model involves a number of variables, one of them being the concentration of dissolved oxygen ($[O_2]$). Considering that the experiments were carried out using vessels always open to air, and given the relatively long time scale of the enzyme inactivation (thousands of seconds), it is assumed that the oxygen concentration is constant during the process and equal to the initial air-saturated value $[O_2]_0$ at the temperature of the experiments, i.e. $[O_2] \approx [O_2]_0 = 0.2$ mM [46]. The following, reasonably simple analytical expression can be derived (see the Supplementary material for the details of the mathematical treatment) [47] to give a good fit to the detected curves [Eq. (1)].

$$[JBU_{act}] = [JBU_{act}]_0 \exp(-k_{app1}t - k_{app2}t^2) \tag{1}$$

In this equation, k_{app1} and k_{app2} are functions of the kinetic constants k_1, k_2, k_3 and k_4 in Scheme 2, as well as the initial concentration of catechol, $[CAT]_0$, as indicated in Eq. (2) and Eq. (3), which were derived through solution of the system of simultaneous ordinary differential equations by applying the steady state approximation and the Taylor series comparison technique [48] (see Supplementary material for the full mathematical treatment):

$$k_{app1} = \frac{k_1 k_3}{k_4}[CAT]_0 \tag{2}$$

$$k_{app2} = \frac{2k_1 k_2 k_3}{k_4}[CAT]_0^2 - \frac{k_1^2 k_3}{k_4}[O_2]_0[CAT]_0 \tag{3}$$

A simultaneous fit of the data obtained for JBU at four different initial catechol concentrations using this model provided a very good agreement between theory and experiment (Fig. 2B), yielding the following values: $k_1 = 0.9 \pm 0.1$ M^{-1} s^{-1}, $k_2 = 49 \pm 4$ M^{-1} s^{-1} and $k_3/k_4 =$

Scheme 2. Working hypothesis for the mechanism of inactivation of urease by catechol.

---

[a] Superoxide is not a reactant in any of the considered reactions in Scheme 2; most probably it will undergo dismutation to generate dioxygen and hydrogen peroxide, but this reaction does not influence the kinetic signature analyzed here.

29 ± 4.[b] The much larger value of k_2 as compared to k_1 is consistent with the observed autocatalytic process, while the k_3/k_4 value significantly larger than one implies that the reaction continues until complete consumption of the active enzyme.

4. Conclusions

The molecular details of the urease irreversible inactivation by catechol have been elucidated. This molecule inactivates urease by a radical-based mechanism that yields a covalent adduct with the conserved αCys322 located on the flap that modulates the opening and closing of the active site channel, blocking it in the open position (see Fig. S6). This seems a general trait, common to aromatic poly-hydroxylated urease inhibitors, as it has been observed recently also for the case of urease irreversible inactivation by *para*-benzoquinone [16]. The role of reducing agents such as sulfite on this reaction has also been clarified. Finally, a kinetic scheme of urease inactivation by catechol was elaborated and positively tested against experimental data. The obtainment of key structure-activity relationships for the development of improved poly-hydroxylated aromatic molecules as urease inactivators, with potential applications in medicine as well as in agro-environmental settings, is under active investigation in our laboratory.

Abbreviations

SPU	*Sporosarcina pasteurii* urease
JBU	*Canavalia ensiformis* (jack bean) urease
CAT	catechol
OBQ	*ortho* benzoquinone
OSQ	*ortho*-semiquinone radical
PBQ	1,4-benzoquinone
BME	2-mercapto-ethanol
DAP	diamidophosphate
DFT	density functional theory
INT	reaction intermediate
TS	transition state

Acknowledgments

Appendix A. Supplementary data

Supplementary data for catechol-bound SPU have been deposited at the Protein Data Bank with accession code 5G4H. Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.jinorgbio.2016.11.016.

[b] This model was not applicable to the case of SPU because the active role of sulfite would impose additional reactions, increasing the number of parameters to fit and decreasing their reliability; furthermore, the use of a much larger concentration of catechol to observe an inhibition effect, due to the presence of sulfite, together with the value of k_1 obtained in the case of JBU, implies that, in the case of SPU, the concentration of dioxygen in solution is no longer constant; these factors would complicate the model for SPU without adding any significant information on the intrinsic inhibition of urease by catechol.

References

[1] B. Zambelli, F. Musiani, S. Benini, S. Ciurli, Chemistry of Ni²⁺ in urease: sensing, trafficking, and catalysis, Acc. Chem. Res. 44 (2011) 520–530.

[2] M.J. Maroney, S. Ciurli, Nonredox nickel enzymes, Chem. Rev. 114 (2014) 4206–4228.

[3] J.C. Rutherford, The emerging role of urease as a general microbial virulence factor, PLoS Pathog. 10 (2014), e1004062.

[4] S. Kiss, M. Simihaian, Improving Efficiency of Urea Fertilizers by Inhibition of Soil Urease Activity, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

[5] J.L. Slonczewski, M. Fujisawa, M. Dopson, T.A. Krulwich, Cytoplasmic pH measurement and homeostasis in bacteria and archaea, Adv. Microb. Physiol. 55 (2009) 1–79 (317).

[6] J.G. Kusters, A.H. van Vliet, E.J. Kuipers, Pathogenesis of *Helicobacter pylori* infection, Clin. Microbiol. Rev. 19 (2006) 449–490.

[7] Schistosomes, liver flukes and Helicobacter pylori, IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, 61, World Health Organization, International Agency for Research on Cancer 1994, pp. 1–241 Lyon, 7–14 June 1994, IARC monographs on the evaluation of carcinogenic risks to humans.

[8] W. Lin, V. Mathys, E.L. Ang, V.H. Koh, J.M. Martinez Gomez, M.L. Ang, S.Z. Zainul Rahim, M.P. Tan, K. Pethe, S. Alonso, Urease activity represents an alternative pathway for *Mycobacterium tuberculosis* nitrogen metabolism, Infect. Immun. 80 (2012) 2771–2779.

[9] G.M. Young, D. Amid, V.L. Miller, A bifunctional urease enhances survival of pathogenic *Yersinia enterocolitica* and *Morganella morganii* at low pH, J. Bacteriol. 178 (1996) 6487–6495.

[10] G.M. Cox, J. Mukherjee, G.T. Cole, A. Casadevall, J.R. Perfect, Urease as a virulence factor in experimental cryptococcosis, Infect. Immun. 68 (2000) 443–448.

[11] B.D. Jones, C.V. Lockatell, D.E. Johnson, J.W. Warren, H.L.T. Mobley, Construction of a urease-negative mutant of *Proteus mirabilis*: analysis of virulence in a mouse model of ascending urinary tract infection, Infect. Immun. 58 (1990) 1120–1123.

[12] P. Kosikowska, L. Berlicki, Urease inhibitors as potential drugs for gastric and urinary tract infections: a patent review, Expert Opin. Ther. Pat. 21 (2011) 945–957.

[13] S. Benini, M. Cianci, L. Mazzei, S. Ciurli, Fluoride inhibition of *Sporosarcina pasteurii* urease: structure and thermodynamics, J. Biol. Inorg. Chem. 19 (2014) 1243–1261.

[14] L. Mazzei, M. Cianci, S. Benini, L. Bertini, F. Musiani, S. Ciurli, Kinetic and structural studies reveal a unique binding mode of sulfite to the nickel center in urease, J. Inorg. Biochem. 154 (2016) 42–49.

[15] S. Benini, W.R. Rypniewski, K.S. Wilson, S. Ciurli, S. Mangani, The complex of *Bacillus pasteurii* urease with β-mercaptoethanol from X-ray data at 1.65 Å resolution, J. Biol. Inorg. Chem. 3 (1998) 268–273.

[16] L. Mazzei, M. Cianci, F. Musiani, S. Ciurli, Inactivation of urease by 1,4-benzoquinone: chemistry at the protein surface, Dalton Trans. 45 (2016) 5455–5459.

[17] J.M. Bremner, L.A. Douglas, Inhibition of urease activity in soils, Soil Biol. Biochem. 3 (1971) 297–307.

[18] R. Puupponen-Pimiä, A.M. Aura, K.M. Oksman-Caldentey, P. Myllärinen, M. Saarela, T. Mattila-Sandholm, K. Poutanen, Development of functional ingredients for gut health, Trends Food Sci. Technol. 13 (2002) 3–11.

[19] E. Pastene, H. Speisky, A. Garcia, J. Moreno, M. Troncoso, G. Figueroa, *In vitro* and *in vivo* effects of apple peel polyphenols against *Helicobacter pylori*, J. Agric. Food Chem. 58 (2010) 7172–7179.

[20] M. Daglia, Polyphenols as antimicrobial agents, Curr. Opin. Biotechnol. 23 (2012) 174–181.

[21] L. Paulo, M. Oleastro, E. Gallardo, J.A. Queiroz, F. Domingues, Anti-*Helicobacter pylori* and urease inhibitory activities of resveratrol and red wine, Food Res. Int. 44 (2011) 964–969.

[22] F. Cardona, C. Andrés-Lacueva, S. Tulipani, F.J. Tinahones, M.I. Queipo-Ortuño, Benefits of polyphenols on gut microbiota and implications in human health, J. Nutr. Biochem. 24 (2013) 1415–1422.

[23] W. Kabsch, Xds, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 125–132.

[24] P. Evans, Scaling and assessment of data quality, Acta Crystallogr. D Biol. Crystallogr. 62 (2006) 72–82.

[25] S. Benini, P. Kosikowska, M. Cianci, L. Mazzei, A.G. Vara, L. Berlicki, S. Ciurli, The crystal structure of *Sporosarcina pasteurii* urease in a complex with citrate provides new hints for inhibitor design, J. Biol. Inorg. Chem. 18 (2013) 391–399.

[26] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, Acta Crystallogr. D Biol. Crystallogr. 53 (1997) 240–255.

[27] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2126–2132.

[28] R.P. Joosten, F. Long, G.N. Murshudov, A. Perrakis, The PDB_REDO server for macromolecular structure model optimization, IUCrJ 1 (2014) 213–220.

[29] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera — a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612.

[30] L. Schrodinger, 2015.

[31] F. Neese, The orca program system, Wiley Interdiscip. Rev.: Comput. Mol. Sci. 2 (2012) 73–78.

[32] A.D. Becke, A new mixing of Hartree–Fock and local density-functional theories, J. Chem. Phys. 98 (1993) 1372–1377.

[33] C. Lee, W. Yang, R.G. Parr, Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density, Phys. Rev. B 37 (1988) 785–789.

[34] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O.

Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M.J. Bearpark, J. Heyd, E.N. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A.P. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N.J. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian, Inc., Wallingford, CT, USA, 2009.

[35] M.J. Frisch, J.A. Pople, J.S. Binkley, Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets, J. Chem. Phys. 80 (1984) 3265–3269.

[36] T.H. Dunning, Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, J. Chem. Phys. 90 (1989) 1007–1023.

[37] R.B. Silverman, Mechanism-based enzyme inactivators, Methods Enzymol. 249 (1995) 240–283.

[38] H. Wagreich, J.M. Nelson, On the oxidation product of catechol when oxidized by means of tyrosinase, J. Biol. Chem. 115 (1936) 459–465.

[39] G.E.K. Branch, M.A. Joslyn, The kinetics of the auto-oxidation of catechol in the presence of several foreign substances, J. Am. Chem. Soc. 57 (1935) 2388–2394.

[40] M.V. Vener, A.V. Odinokov, C. Wehmeyer, D. Sebastiani, The structure and IR signatures of the arginine-glutamate salt bridge. Insights from the classical MD simulations, J. Chem. Phys. 142 (2015) (215106).

[41] M. Kot, W. Zaborska, Irreversible inhibition of jack bean urease by pyrocatechol, J. Enzyme Inhib. Med. Chem. 18 (2003) 413–417.

[42] S. Benini, W.R. Rypniewski, K.S. Wilson, S. Miletti, S. Ciurli, S. Mangani, A new proposal for urease mechanism based on the crystal structures of the native and inhibited enzyme from *Bacillus pasteurii*: why urea hydrolysis costs two nickels, Structure 7 (1999) 205–216.

[43] P. Politzer, J.S. Murray, P. Lane, σ-Hole bonding and hydrogen bonding: competitive interactions, Int. J. Quantum Chem. 107 (2007) 3046–3052.

[44] X.M. Shen, G. Dryhurst, Iron- and manganese-catalyzed autoxidation of dopamine in the presence of L-cysteine: possible insights into iron- and manganese-mediated dopaminergic neurotoxicity, Chem. Res. Toxicol. 11 (1998) 824–837.

[45] N. Schweigert, A.J. Zehnder, R.I. Eggen, Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals, Environ. Microbiol. 3 (2001) 81–91.

[46] R.F. Weiss, The solubility of nitrogen, oxygen and argon in water and seawater, Deep-Sea Res. 17 (1970) 721–735.

[47] G. Lente, Deterministic Kinetics in Chemistry and Systems Biology, Springer International Publishing, 2015.

[48] Á. Balogh, G. Lente, J. Kalmár, I. Fábián, Reaction schemes that are easily confused with a reversible first-order reaction, In. J. Chem. Kinet. 47 (2015) 773–782.