**Thirty Years in the Design and Development of Cruzain Inhibitors**

<table>
<thead>
<tr>
<th>Journal:</th>
<th><em>Journal of the Brazilian Chemical Society</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>JBCHS-2024-0025.R2</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Review</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>03-May-2024</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Jasinski, Gabriel; Universidad de Buenos Aires Martini, M. Florencia; Instituto Tecnológico de Buenos Aires Moglioni, Albertina; Universidad de Buenos Aires, Instituto de Química y Metabolismo delFármaco (IQUIMEFA); Universidad de Buenos Aires Facultad de Farmacia y Bioquímica,</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Chagas disease, discover and development, cruzipain/cruzain inhibitors classification</td>
</tr>
<tr>
<td>Please, specify if the submission is for a Regular Issue or Special Issue:</td>
<td>MedChem - Special Issue</td>
</tr>
</tbody>
</table>

https://mc04.manuscriptcentral.com/jbchs-scielo
This review presents the main developments in inhibitors of cruzipain/cruzain, a crucial enzyme in the life cycle of *Trypanosoma cruzi*, the etiological agent of Chagas disease, from the 1990s to the present.

This work is dedicated to Prof. Eliezer J. Barreiro in recognition of his hard work in medicinal chemistry, which has contributed to inspiring researchers, professionals, and new generations of medicinal chemists.
Thirty Years in the Design and Development of Cruzain Inhibitors

Jasinski Gabriel\textsuperscript{1,2}, Martini M. Florencia\textsuperscript{3,4}, Moglioni Albertina*\textsuperscript{1,2}

\textsuperscript{1}Universidad de Buenos Aires, Facultad de Farmacia y Bioquímica, Cátedra de Química Medicinal, Buenos Aires, C1113AAD, Argentina

\textsuperscript{2}CONICET-Universidad de Buenos Aires, Instituto de la Química y el Metabolismo del Fármaco (IQUIMEFA), Buenos Aires, C1113AAD, Argentina

\textsuperscript{3}Instituto Tecnológico de Buenos Aires (ITBA), Buenos Aires, C1437ETC, Argentina.

\textsuperscript{4}Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, C1437ETC, Argentina.

Corresponding author: *Moglioni Albertina. Química Medicinal, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires. e-mail: amoglio@ffyb.uba.ar https://orcid.org/0000-0002-5625-3627
Cruzipain is the principal protease of *Trypanosoma cruzi*, the etiological agent of Chagas disease. Since its discovery in the 1980s and the resolution of the crystal structure of cruzain (a truncated recombinant form of the enzyme) in 1995, this target has attracted the interest of many research groups for screening studies, structure-based and ligand-based drug design campaigns, which include peptide-like and non-peptide synthetic compounds. In this context, empirical and computational methods have proven to be valuable tools for the study of mechanisms of action, potential binding modes and structure-activity relationships for a diverse series of cruzain/cruzipain inhibitors. This paper, therefore, reviews some of the most relevant chemical groups reported as cruzain inhibitors over the last 30 years of research.

**Keywords:** Chagas disease, discover and development, cruzipain/cruzain inhibitors classification

**Introduction**

Since its discovery in 1909 as an infectious disease caused by *T. cruzi*, Chagas disease has been an object of interest in medicinal chemistry in the search for compounds with both *in vitro* and *in vivo* trypanocidal activity. In this context, the first screening experiments in mice and guinea pigs infected with *T. cruzi* were conducted between 1912 and 1914 by Mayer and Rocha-Lima in laboratories of the current Bernhard-Nocht-Institut für Tropenmedizin in Hamburg, Germany.\(^1\)\(^-\)\(^3\) These experiments, which included compounds used to treat other tropical diseases, showed that *T. cruzi* was remarkably resistant to large and diverse series of compounds (as confirmed by subsequent screening experiments).\(^4\) It was not until the second half of the 1930s, with the exploration of aminoquinaldines and arsenical derivatives of benzisoxazole at Bayer/IG Farben AG laboratories that the first active compounds were found in animal models of infection. Between 1937 and 1953, the
clinical study of the compounds BAY-7602, BAY-9736 and BAY-10557 (Spirotrypan) (Figure 1.A, 1, 2 and 3, respectively) was carried out in acute patients under the supervision of Salvador Mazza and Cecilio Romaña in Argentina, observing that they were active in vivo against the blood forms of the parasite (trypomastigotes). However, the lack of activity against the intracellular forms (amastigotes) and the development of toxicity in patients led to relapses and undesirable side effects, precluding the success of the experimental therapy. 5–8

---

**Figure 1.** A. Structure of the first compounds used in clinical trials on patients with Chagas’ disease. B. Route of development of nifurtimox and benznidazole as anti-T. cruzi drugs. C. Other compounds with trypanocidal effects on T. cruzi.
Although some new compounds allowed to obtain interesting results, such as those corresponding to Carbidium/BW-74C48 (a phenanthridinium derivative) (Figure 1.A, 4), developed by Wellcome Laboratories,9 the observation of the trypanocidal action of nitrofurazone 5 in 1952 by Ardzoony Packchanian of the University of Texas10 represents a fundamental discovery, being the first case of a compound capable of reaching a level of activity sufficient to be applicable in the therapy of the disease. Although concerns raised about its toxicity prevented its introduction into clinical application for Chagas disease treatment (provided that nitrofurazone clinical trials on patients infected with Trypanosoma brucei had to be discontinued in 1960 due to an unacceptable security profile),11 this compound provided a basis for the development of more potent and safer analogues. In this context, two types of nitro heterocyclic compounds were developed: 5-nitrofurfural 6 and 2-nitroimidazole derivatives 7-12 (Figure 1.B-C).

The design of 5-nitrofurfural derivatives, started by Bayer in 1962, led to nifurtimox 6,12,13 the first widely used drug for the treatment of acute Chagas disease. In parallel, the development of 2-nitroimidazole derivatives by F. Hoffmann-La Roche laboratories incorporated elements of the metronidazole (9) design route (which emerged in 1959 from the screening of more than 200 analogues of azomycin14), leading to the development of benznidazole (Figure 1.B, 7),15,16 still used as a first-line treatment for Chagas disease. The toxicity problems observed in therapeutics based on this class of nitro compounds can be explained by the low selectivity of its mechanism of action, based on the generation of intense radical stress that can affect both the trypanosome and the host cells.
Although nifurtimox 6 and benznidazole 7, have been available for the treatment of Chagas disease since 1972 as drugs whose indication is limited to the acute phase of the disease, their adverse effects led to the persistent search for safer and more effective compounds capable of exploiting the biomolecular, cytoarchitectonic and metabolic differences between the parasite and host cells, to obtain maximum efficacy, selectivity, and safety in pharmacological treatments. In this context, the study of the comparative molecular biology of *T. cruzi* and its mammalian hosts has enabled the identification and validation of several biomolecular targets to guide drug design.

According to the classification of *T. cruzi* targets for trypanocidal drugs in three groups proposed by Duschak in 201917 (group I: main molecular targets such as specific enzymes involved in essential processes for parasite survival; group II: biological pathways and their key specific enzymes, and Group III: atypical organelles/structures present in the parasite relevant clinical forms), the proteases (included in group I) represent one of the most relevant biomolecular drug targets considered for the development of new treatments for Chagas disease. In general terms, the interest in the development of protease inhibitors is motivated by the diversity and significance of the processes in which this class of enzymes participate, including digestion of exogenous proteins, regulation of the half-life of protein components intracellular and regulation of processes by controlled proteolysis of other enzymes and protein components.18

In the protease classification system, taxonomy divides these enzymes into families (proteases homologous to each other, with closely related sequences) and clans (groups of evolutionarily related families). In the case of *T. cruzi*, its genome encodes almost 400 putative proteases, including nearly 70 cysteine proteases, 40 serine proteases, 250 metalloproteases, 25 threonine proteases and 2 aspartyl proteases.19 Such diversity reflects the importance of proteases in the biology of *T. cruzi*. In this
context, cruzipain (CZP, EC: 3.4.22.51) is one of the most exhaustively studied *T. cruzi* proteases. This enzyme is a highly glycosylated cysteine protease, identified by Cazzulo et al. in 1989. Being the most abundant protease of the parasite, it is expressed in all stages of its development and participates in fundamental processes (Figure 2), such as:

a) degradation of proteins introduced into the epimastigote reservoir,

b) metacyclogenesis (conversion of epimastigotes into metacyclic trypomastigotes),

c) evasion of the host immune system by antibody proteolysis,

d) reduction of the antiparasitic response capacity of macrophages, activation of TGF-β, increase of the arginase activity and degradation of the transcription factor NF-κB, which facilitates the persistence of the infection,

e) increase of infectivity by activation of the kinin-mediated signalling pathway, and

f) degradation of components of the extracellular matrix (fibronectin and collagens type I and IV).
Figure 2. CZP functions in the context of host cell- \textit{T. cruzi} interactions. A and B. Interactions between CZP and immune system elements and non-phagocytic cells in mammalian hosts. The red arrows indicate unfavoured or deactivated processes or elements, while green arrows indicate favoured processes in presence of CZP.

Given the multiple functions of CZP and its importance for the survival, development, and virulence of \textit{T. cruzi}, it is not surprising that potent inhibition of this cysteine protease implies devastating consequences for the parasite life cycle. These effects have been demonstrated in \textit{in vitro}
experiments,\textsuperscript{33–35} and in animal models of Chagas disease.\textsuperscript{36} Therefore, CZP constitutes a validated target for the design of antichagasic drugs.

**Cruzipain/cruzain as a molecular target for drug development**

CZP is encoded in the *T. cruzi* genome in genes of 1407 base pairs located in tandem repeat copies [with up to 130 total copies per genome in the case of the Tulahuen-2 (Tul2) strain]\textsuperscript{37} at syntenic loci distributed on several chromosomes, where there is polymorphism between gene copies (Figure 3).\textsuperscript{38} The tandem arrangement includes possible polymorphic variations between copies of the CZP gene, each with potentially different drug sensitivity.

![Trypanosoma cruzi genome](https://mc04.manuscriptcentral.com/jbchs-scielo)

**Figure 3.** Localization of a CZP gene tandem in the *T. cruzi* genome, gene expression and enzyme maturation. The primed and unprimed residue numbering refer to the pre-proenzyme and the mature enzyme, respectively. The catalytic domain corresponds to the sequence of the cruzipain, as shown.

https://mc04.manuscriptcentral.com/jbchs-scielo
As is usual in trypanosomatids, the regulation of CZP gene expression is post-transcriptional. Thus, although the mRNA levels of the enzyme remain stable throughout the parasite life cycle, the levels of CZP synthesised strongly depend on the considered stage. In this context, the expression level of CZP in trypomastigotes and amastigotes is approximately 20-25% of that corresponding to epimastigotes. It has been observed that the subcellular localization of the enzyme is stage-dependent. Therefore, while CZP is concentrated in the reservosomes of epimastigotes, it is located in the region of the flagellar pocket of trypomastigotes and on the surface of the amastigotes of T. cruzi (in contact with the host cell cytoplasm). This implies different acidic microenvironments for CZP since the reservosomes in epimastigotes and the parasite surface in amastigotes have different pH values (6.3 and 7.4, respectively), which is an element to consider in the design of inhibitors.

The enzyme is synthesised in the endoplasmic reticulum (ER) of the parasite as a pre-proenzyme, which consists of a signal peptide, a N-terminal domain, a catalytic domain, and a long C-terminal domain of unusual length among proteases. During the maturation of the enzyme, cleavage and elimination of the N-terminal domain occurs. While in transit through the ER and the Golgi system, some asparagine residues in CZP (mainly Asn255) are modified by anchoring high mannose-type, hybrid monoantennary or complex biantennary oligosaccharides with diverse levels of sulfation. This, added to the observed gene polymorphism, clearly shows that CZP is a complex and heterogeneous enzyme.

As a validated target for anti-Chagas drug’s design, CZP has been extensively studied using X-ray diffraction techniques from a recombinant form known as cruzain (CZ). The latter results from the loss of the C-terminal domain of CZP, which has allowed its crystallization. These
experiments provided detailed structural information about CZ complexes with different inhibitors, whose structures are available in the PDB database. Given the available information, it is possible to represent the tertiary structure of this enzyme and its catalytic domain, as illustrated in Figure 4.

**Figure 4.** A. 3D structure of CZ (PDB ID: 3LXS) and details of its active site, with catalytic residues indicated in red. The original ligand was removed. B. Enzyme surface indicating the localization of the subsites S3 to S2′ and catalytic residues C25 and H162. Images rendered with UCSF Chimera.

The classification of CZP into the C1 family of proteases (which includes human cathepsin L as one of its members) implies the definition of its catalytic site in terms of two groups of residues:

a) residues directly involved in the catalytic action (Cys25 and His162), and

b) residues whose influence increases the enzyme reactivity (mainly Asn182, with possible participation of Gln19).

Several mechanisms have been proposed in the literature to describe the corresponding catalytic cycle, one of which is illustrated in Figure 5.
Figure 5. Catalytic cycle of CZ/CZP for cleavage of a peptidic substrate, adapted from the mechanism postulated in Ref. 49. P1 and P1’ refer to residues that binds to S1 and S1’ sites on the surface of the protease, respectively. The catalytic (C25 and H162) and auxiliar residues (Q19 and N182) are indicated by red and grey boxes, respectively.

The mechanism of action of CZ has been studied from both experimental (through kinetic experiments and directed mutagenesis)\textsuperscript{49} and computational points of view (through the use of QM/MM methods),\textsuperscript{50–52} giving rise to controversies about:

a) the precise sequence of steps throughout the catalytic cycle,

b) the formation and reactivity characteristics of the thiolate/imidazolium pair; and

c) the role of residues close to the Cys25-His162 (red boxes in Figure 5) dyad in the catalytic mechanism.
From the studies developed it can be concluded that (see Figure 5):

1) The catalytic cycle develops in two stages:
   a) binding of the substrate to the catalytic site and formation of the S-acylated intermediate III (by the nucleophilic attack of the thiolate form of Cys25 on the carbonyl group of the peptide bond of the substrate),
   b) deacylation of Cys25 by nucleophilic attack of water assisted by His162 (intermediate IV), completing the catalytic cycle through intermediates V and VI, with release of the corresponding products of peptide cleavage.

2) The formation of the thiolate/imidazolium pair (a fundamental step for most of the kinetic models developed for this class of proteases) seems to be influenced by the presence of the substrate in the active site. Therefore, the reactivity of this pair could be modulated by elements such as the spatial orientation of the imidazole ring of His162 under the influence of Asn182 (although not considered an essential residue for the catalytic action, the replacement of Asn182 (grey box in Figure 5) by other amino acids generates a significant reduction in enzymatic activity as seen in directed mutagenesis experiments carried out on papain53),

3) The role of Gln19 (grey box in Figure 5) seems to be related to the definition of an ‘oxyanion hole’ that stabilises transition states involved in the formation of tetrahedral intermediates produced as a consequence of nucleophilic attacks on the C(sp²) centres in intermediates I and IV of the catalytic cycle. Results obtained in directed mutagenesis experiments on papain suggest that Gln19 seems not to be essential for enzymatic activity; however, its replacement by other amino acids significantly reduces the hydrolysis rate of peptide substrates.
4) The results obtained in kinetic experiments\textsuperscript{49,54} and QM/MM simulations\textsuperscript{52} suggest that the deacylation of the intermediate III is the rate-limiting step. In particular, the QM/MM study of the Gibbs free-energy profile associated with the catalytic cycle allows to infer the existence of a deep energy minimum between the acylation and deacylation stages, which corresponds to a state that involves the $S$-acylated intermediate (III). The relative stability of this intermediate makes it relatively difficult to proceed to the deacylation stage through intermediate IV or return to the tetrahedral intermediate II.

According to the convention adopted by Schechter and Berger\textsuperscript{55} to identify the residues of the substrate and the subsites of the enzyme, ‘Pn’ is designated to the residues present in the substrate and ‘Sn’ refers to the corresponding subsites of the enzyme. The residues identified as Pn' are located towards the C-terminal of the peptide bond, while the residues Pn are located at the N-terminal. Thus, residues P3, P2, P1, P1' and P2' are located in the enzymatic subsites S3, S2, S1, S1' and S2', respectively. The arrangement of the ‘Sn’ subsites on the CZ surface is indicated in Figure 4.B.

As a member of the C1 family of cysteine proteases, the specificity of CZP mainly depends on the properties of the S2 sub-site, which have in this case the following characteristics:

a) it is a relatively wide and deep site,

b) it is composed mostly of hydrophobic residues, and

c) it has an ionizable and flexible residue at its distal end (Glu208).

Therefore, the S2 sub-site confers selectivity towards substrates that have hydrophobic residues (mainly Leu, Phe and Val) or polar residues (Arg and Tyr) at the P2 location.\textsuperscript{56,57} In this
context is remarkable the role of Glu208, capable of swinging into the S2 sub-site to establish electrostatic or hydrogen bonding interactions with polar residues at P2 or swing out of the S2 sub-site allowing the entry of hydrophobic residues into S2.\textsuperscript{46} In parallel, the S1 and S1' sites confer some selectivity toward substrates with basic residues at the P1 position (e.g. Lys and Arg) and hydrophobic residues at the P1' position (e.g. Leu and Phe), respectively.\textsuperscript{57}

Comparative studies between the kinetic profiles corresponding to natural CZP (extracted from epimastigotes of a virulent strain of \textit{T. cruzi}) and recombinant CZ indicate the presence of some differences related to kinetic constants and thermodynamic parameters (such as activation energy, activation entropies and affinity for synthetic substrates like Abz-KLRFSKQ-EDDnp).\textsuperscript{58} However, most molecular studies of potential inhibitors have been carried out against CZ instead of CZP, the latter being present in \textit{T. cruzi} parasites as a polymorphic mixture containing CZ as one of its possible variants. This situation can be explained by the lack of availability of reliable 3D models of the CZP structure for computational studies and the advantages of having a recombinant enzyme that is homogeneous in composition (without the complexity represented by natural CZP).

The enzymatic activity of CZP in \textit{T. cruzi} is regulated by a series of mechanisms that include specific inhibitors such as chagasin,\textsuperscript{59-61} a protein functionally related to the cystatin family. The importance of chagasin in the control of CZP activity and its high expression levels in the parasite make it a possible target for the development of experimental vaccines, as demonstrated by Malchiodi \textit{et al.}\textsuperscript{62}
Development of cruzipain/cruzaín inhibitors

Validated as a key biomolecular drug target,\textsuperscript{42,63} CZ/CZP has been at the core of intense investigations in search of potent and effective inhibitors with potential clinical applications, which represents a challenging task, provided that this enzyme belongs to the same cysteine protease family that includes human cathepsins (e.g. CatK and CatL), which impose selectivity as an important requirement. In this sense, it is worth highlighting the works of Montanari \textit{et al.}\textsuperscript{42,64,65} on the development of QSAR models and molecular dynamics-based techniques with good predictive value, capable of classify ligands in terms of their potential selectivity towards CZ or cathepsins.

Of note, in the absence of empirical evidence about the corresponding type of inhibition, it is frequent to implicitly assume that most of the inhibitors act through competitive mechanisms (binding to sites near or at the catalytic site), regardless of whether they are peptidyl compounds or not. However, it is necessary to consider the possibility of alternative mechanisms that involve binding to potential allosteric sites, as suggested by Hernández Alvarez \textit{et al.}\textsuperscript{66} based on molecular dynamics studies.

The development of CZ inhibitors has given rise to a great diversity of compounds that could be classified into two general categories: those of peptide nature and those of non-peptide nature.\textsuperscript{21} Within both groups, inhibition can be reversible or irreversible.

In the following sections, the most representative peptide and non-peptide inhibitors described in the literature to date will be reviewed, as well as those derived from drug repositioning or natural sources.
1. Peptide inhibitors

This group includes peptidomimetic type inhibitors that have natural and/or modified amino acids in their structure. Analysis of the molecular architecture of this class of inhibitors indicates that they have three structural regions that determine their activity, as indicated in Figure 6.A.

**Figure 6.** A. Typical structure of peptide type inhibitors of CZ, represented by K11777. Terminal, recognition and reactive ("warhead") groups are indicated in green, blue, and red, respectively. B. Proposed action mechanism of K11777 on CZ, adapted from Ref. 79.

These inhibitors generally contain an electrophilic group (e.g. a vinyl sulfone, Figure 6.B) capable of reacting with the thiolate/thiol group of the catalytic residue Cys25 through a nucleophilic attack, establishing a covalent bond between the enzyme and the inhibitor. Depending on the identity of the reactive group and the stability of the bond established between the enzyme and the ligand, the inhibitory effect can be reversible or irreversible. In parallel, the presence of the recognition group is essential not only to affinity but also to establish the selectivity of the action.
1.1. Irreversible peptide inhibitors

The action of this class of inhibitors is due to the presence of reactive groups such as α-diazoethyl ketone, α,β-epoxy ketone, α-halomethyl ketone, allyl sulfone and vinyl sulfone (Figure 7, 13-18), all of them capable of producing an irreversible alkylation of the thiol/thiolate group present at the catalytic residue Cys25.

![Chemical Structures](image)

**Figure 7.** Irreversible peptide-like inhibitors of CZ/CZP. Terminal, recognition and reactive (“warhead”) groups are indicated in green, blue, and red, respectively.

The potency of this class of compounds as *in vitro* CZ inhibitors is generally in the nM-μM range, with a time-dependent action profile which is typical for ligands whose mechanism involves the formation of covalent bonds. The interest on the definition of the binding mode and mechanism of action of these compounds has led to the resolution of the structure of a series of CZ-inhibitor complexes by X-ray crystallography (e.g. Z-Phe-Ala-FMK, K11777 and WRR-483) and the application of computational tools for simulation of enzyme-inhibitor reactions.
Taking the inhibitors of the vinyl sulfone family as an example of compounds capable of acting as Michael acceptors for the nucleophilic attack of Cys25, the evidence indicates that this reaction progresses through a concerted mechanism, where the nucleophilic attack and the capture of a proton from the imidazolium ring of His162 proceed simultaneously (Figure 6.B).\textsuperscript{79}

In this context, one of the most promising compounds has been K11777 \textbf{13}, developed by McKerrow’s group, with powerful and effective action on \textit{T. cruzi}, both \textit{in vitro}\textsuperscript{63,75} and \textit{in vivo} (considering murine\textsuperscript{36,80} and canine models\textsuperscript{81}). However, advanced preclinical tests indicated tolerance problems in dogs and primates, leading to the interruption of the development process of K11777 as a possible drug for treatment of Chagas disease.\textsuperscript{82} Complementary studies have identified unfavourable pharmacokinetic characteristics of this compound, such as cytochrome inhibition (e.g. CYP3A4)\textsuperscript{83} and P-glycoprotein-dependent oral absorption.\textsuperscript{84a}

The preference for hydrophobic or polar residues at the P2 position in CZ substrates led to an investigation of the effect of replacing the Phe residue at K11777 (\textbf{13}) with a hydrophilic residue such as Arg on inhibitory activity. The new ligand (WRR-483 \textbf{15}, Figure 7) not only showed potent and pH-dependent activity in CZ assays (with pIC\textsubscript{50} values of 70 nM and 8 nM at pH 5.5 and 8.0, respectively) but also a marked anti-\textit{T. cruzi} effect in both \textit{in vitro} and \textit{in vivo} models.\textsuperscript{48} However, the low oral bioavailability of this compound indicated the need for further structural modifications.

Although the precise pattern of non-covalent interactions between this type of inhibitors and CZ/CZP could be optimised to achieve a more selective profile of action, the potential irreversible
off-target effects remain a serious concern in the development of new drugs for Chagas disease treatment.\textsuperscript{84b}

1.2. Reversible peptide inhibitors

As was mentioned above, the development of enzyme inhibitors with irreversible action is limited. For this reason, the development of new drugs directed towards CZ/CZP has incorporated the design of peptide-like inhibitors with a reversible effect.\textsuperscript{84b} Some of the most important developments are those related to the discovery of a) the vinyl sulfone WRR-669 \textsuperscript{22}, b) the dipeptidyl nitriles Cz-007 \textsuperscript{24} and Cz-008 \textsuperscript{25}, and c) the α-ketoamide AQ903084 \textsuperscript{26} and the α-ketoester AQ581332 \textsuperscript{28}.

a. Vinyl sulfone WRR-669

The presence of unfavourable characteristics in vinyl sulfones with irreversible action, such as K11777 \textsuperscript{13} and WRR-483 \textsuperscript{15}, led to the design of new analogues. In this context, in an attempt to increase the oral bioavailability of WRR-483 \textsuperscript{15}, the methylene-guanidine moiety at the Arg residue was replaced by an oxyguanidine group (giving rise to WRR-662 (Figure 8, compound \textsuperscript{19a})). The rationale behind this modification (which involved the replacement of Arg by a canavanine residue, an unnatural amino acid) was that a significant reduction in the pKa value of the residue at the P2 position (from 12.48 for Arg to 7.01 for canavanine)\textsuperscript{85a} could provide an increase of the unionized fraction of the ligand at physiological pH and a correspondingly improved oral bioavailability.
Additional structural modifications, such as the homologation of the alkyl side chain of the canavanine residue and the replacement of the phenyl ring at the vinyl sulfone warhead by a 2-pyrimidinyl group, conduct to the analogue WRR-669 (22), a compound with a surprising dual mechanism of action: a time-dependent action \textit{in vitro}, achieving the inhibitory effect through the formation of a covalent bond with CZ under acidic conditions (pH = 5.5), and a reversible inhibition of CZ without covalent binding at pH = 8.0 (inferred by analysis of kinetic data and confirmed by X-ray diffraction studies of WRR-669-CZ crystals).\textsuperscript{85b}
Figure 8. Discovery route of WRR-669. The values in brackets represent the relative potency of the inhibitors based on the corresponding $k_{inact}/K_i$ value for WRR-483 considering pH = 5.5.

b. Dipeptidyl nitriles

The reversible nature of the formation of adducts (thioimidates) obtained by reaction between thiolate and nitrile groups (Figure 9. A)\[^{86-89}\] suggests that the incorporation of the latter into cysteine
protease inhibitors as an electrophilic warhead could represent an alternative in the design of new CZ inhibitors. In view of the similarities between CZ and human cathepsins, a library of nitrile compounds related to odanacatib (23) (developed by Merck as a cathepsin K inhibitor for the treatment of osteoporosis) was screened against CZ. This led to the identification of two odanacatib analogues with IC$_{50}$ values in the nM range (compounds 24 and 25a). The systematic structural modification of the terminal and recognition groups of 25a provided a series of new inhibitors of CZ, of which 25b represents an optimised example in terms of potency. Although subsequent testing of 24 and 25b in murine models of T. cruzi infection have demonstrated high cure rates (70% and 90%, respectively), the relatively low selectivity and modest pharmacokinetic properties (with oral bioavailability close to 45-50% and plasma half-lives between 1 and 4 h) indicate the need for further structural optimisation to enable their development as drugs for the treatment of Chagas disease.90

Figure 9. A. Mechanism of action proposed for reversible nitrile inhibitors, adapted from reference 87. “TG” and “RGs” represent terminal and recognition groups, respectively. B. Representative examples of nitrile-type inhibitors of CZ. The electrophilic warhead, and the recognition and terminal groups has been indicted in red, blue, and green, respectively.
c. α-ketoester AQ581332 and α-ketoamide AQ903084

Between 1975 and 1993, the study of the mechanism of action of peptidyl-α-keto acid/ester/amide as reversible inhibitors of proteases,91,92 led to the development of reversible inhibitors of cysteine proteases such as cathepsin B and papain.93a As proposed, the mechanism of action for these compounds involves the nucleophilic attack of the catalytic Cys residue on the corresponding α-keto group.93a Although this first generation of compounds had low potency (with \( k_i \) values of the order \( \mu M \)), their discovery stimulated its development as cysteine protease inhibitors. Thus, the analysis of a library of peptidyl-α-keto amides in search of CZ inhibitors allowed Craik et al. to find compounds capable of inhibiting this enzyme at the nM level.93b Systematic modifications conducted to an extensive series of analogues that includes compounds 26-28 (Figure 10), with activity on CZ at the nM level (e.g., vinyl-α-ketoester AQ581332) and sub-nM (e.g. α-ketoamide AQ903084).93b

![Figure 10](https://mc04.manuscriptcentral.com/jbchs-scielo)

**Figure 10.** Reversible peptide inhibitors of CZ belonging to α-ketoester and α-ketoamide families. The arrows indicate the site for nucleophilic attack by the catalytic C25 residue.

Unfortunately, these analogues have demonstrated limited *in vivo* activity, mainly attributed to difficulties in penetrating cell membranes, with a consequent reduction in the bioavailability of the compound for effective blocking of the CZP enzyme of the parasite.93b
2. Non-peptide inhibitors

The limitations found in the design of peptide inhibitors are frequently associated with inadequate pharmacokinetics (e.g. low oral bioavailability and reduced plasma half-life). The presence of peptide fragments represents both an advantage and a potential weakness. Although peptide fragments could favour the action on the target given some local molecular similarity with the natural substrates for an enzyme, they also facilitate its recognition and clearance by metabolic processes. Likewise, it hinders its passage through biological membranes and access to biomolecular targets located in some pharmacokinetic compartments. In this context, the development of non-peptide inhibitors allows greater versatility in design and potential access to a better pharmacokinetic profile.

As observed in the literature, the non-peptide inhibitors represent a large and structurally diverse group, including both reversible and irreversible inhibitors, with some of them reaching activity values in the nM order.

2.1. Irreversible non-peptide inhibitors

Research on this class of inhibitors has given rise to various families of compounds, among which stand out those derived from α-(aryloxy)methyl ketones and α-(acyloxy)methyl ketones. As in the case of peptide-type inhibitors, there is a group that allows the recognition of these by CZ, and an electrophilic warhead capable of reacting with the thiolate/thiol group of the catalytic residue Cys25, leading to its irreversible S-alkylation (Figure 11.A).
Based on the application of Cu(I)-catalysed azide-alkyne cycloaddition\(^6\) for the fast construction of combinatorial libraries of 1,2,3-triazole derivatives joining molecular fragments by click chemistry reactions and the use of a search strategy known as Substrate Activity Screening (SAS), Ellman et al.\(^8\) found a series of non-peptide structural fragments capable of binding CZ with high affinity (recognition groups). Taking into account what was found by Krantz et al.\(^9\) regarding the use of reactive fragments of the \(\alpha\)-(phenoxy)methyl ketone and \(\alpha\)-(benzoyloxy)methyl ketone type, Ellman et al.\(^8\) identified the groups \(\alpha\)-(2,3,5,6-tetrafluorophenoxy)methyl ketone and \(\alpha\)-[2,6-bis(trifluoromethyl) benzoyle]methyl ketone as fragments that, combined with the recognition groups, were capable of give ligands with both potent and irreversible inhibitory action on CZ (Figure 11.B).

The study of these derivatives allowed the observation of potent effects in tests on infected macrophages (IC\(_{50}<10\ \mu\text{M}\)) without developing signs of cytotoxicity (CC\(_{50}\geq10\ \mu\text{M}\)).\(^5\) The ability of one of the compounds (KB2, 29) to eradicate the parasite in a murine model of the disease without
apparent toxicity to the host and the subsequent discovery of analogues with greater potency (e.g. 30, with IC$_{50}$ values of 0.003 μM both against CZ and T. cruzi amastigotes, respectively), are encouraging results for the continued development of this class of derivatives as potential antichagasic drugs.

2.2. Reversible non-peptide inhibitors

Starting with the pioneering work of Cohen et al. on the search for reversible non-peptide inhibitors of CZ, based on the enzyme crystallographic structure and molecular docking, the development of this type of ligands has led to a wide diversity of structures. The most relevant of which are described below.

a) N-acyl hydrazones (NAHs)

The NAHs can be considered as aza-vinylogues of the peptide scaffold (Figure 12.A) where the replacement of the peptide backbone provides a more rigid central structure, providing an interesting strategy for the design of protease inhibitors.

The origin of this family is related to the search for non-peptide inhibitors of falcipain-1, based on the use of aromatic acyl hydrazides. Cohen et al. expanded the search to CZ inhibitors through the combined use of in silico search and enzyme inhibition experiments, finding compounds with IC$_{50}$ values in the μM range. Subsequently, the analysis of a new series of 112 analogues allowed the discovery of compounds of greater potency and the identification of relevant physicochemical
properties through Quantitative Structure-Activity Relationship (QSAR) studies: Comparative Molecular Field Analysis (CoMFA) and Hologram QSAR (HQSAR) (Figure 12.B, compounds 32 and 33). In the case of 33 (Figure 12.B), an inspection of the steric contour plots of CoMFA studies revealed that the sterically favourable regions were located near the pyrazole group. The electrostatic contribution contour surrounding the 4-position in the thiophene ring show that negative potentials in this position are likely to increase the inhibitory profile of the compound. However, the strong steric hindrance at this position makes this site a chemically a challenging site for substitution. The predictive HQSAR models, for these studied compounds, are readily derived using only atoms, bond, and connectivity-distinction information. Adding other fragment distinction into molecular hologram does not appear to improve the model as measured by statistical parameters ($q^2$ and $r^2$).

The discovery of the ability of furoxane compounds to promote the release of nitric oxide (NO) led to the incorporation of this heterocyclic group in acyl hydrazide-type structures, looking for compounds with a dual mechanism of action. The tests conducted on the series of furoxan derivatives did not show a linear correlation between the NO release capacity and the deleterious effect on T. cruzi. However, it was possible to find compounds active on CZ at the µM level, capable of acting as trypanocides in similar concentration levels (Figure 12.B, compounds 34 and 35) through mechanisms that potentially involve diverse molecular targets in addition to CZP.
Figure 12. A. NAHs as aza-vinylogues of the peptide scaffold. B. Examples of NAH inhibitors of CZ. The N-acetylhydrazone group was indicated in blue.

b) Thiosemicarbazones (TSCs)

TSCs have been of interest to medicinal chemistry for more than 70 years.\textsuperscript{107-116} One of the most relevant applications of TSCs is related to the design of compounds directed against parasites of the order Kinetoplastida, giving rise to compounds with trypanocidal effects,\textsuperscript{90,117} some of which are mediated by CZ/CZP inhibition, as described below.

In particular, the interest in TSCs as compounds with action on \textit{T. cruzi} began with the discovery of aryl-aldehyde derivatives, active in both \textit{in vitro} and \textit{in vivo} models, by H. Robert Wilson in 1974\textsuperscript{118} (Figure 13.A). The discovery of more than 30 TSCs with inhibitory effects on CZ by McKerrow \textit{et al.} in 2002 (as part of a collaborative work with Parke-Davis)\textsuperscript{119} represents a landmark in the development of reversible non-peptidic inhibitors for CZ (Figure 13.C). These TSCs 38-43, derived from aryl-aldehydes and alkyl-aryl-ketones, were notable for their high to moderate inhibitory potency on CZ (with IC\textsubscript{50} values between 10 \(\mu\text{M}\) and 0.020 \(\mu\text{M}\)), the trypanocidal effect
of some members of the series, and the ability to act on homologous enzymes in other Kinetoplastida (such as *T. brucei* and *Leishmania sp.*). As will be shown later, these TSCs represented the starting point for the application of pharmacomodulative strategies that gave rise to derivatives with different degrees of inhibitory efficacy on CZ and parasites.

**Figure 13.** A. Trypanocide TSCs identified by H.R. Wilson in 1974. B. Mechanism of action proposed for TSC as reversible inhibitors of CZ, adapted from Ref. 119. C. Inhibitors of CZ developed by McKerrow et al.119.

Given the time-dependent inhibition observed for a series of TSCs, some authors have suggested that the mechanism behind the inhibitory effect on CZ involves the formation of covalent bonds as a product of the interaction of these ligands with the catalytic dyad Cys25-His162. (Figure 13.C) In this context, the analysis of the spatial disposition adopted by the TSC group relative to the
location of Cys25 and His162 residues in molecular docking experiments on the structure of CZ suggests the feasibility of the proposed mechanism.\textsuperscript{119} Later, Trossini \textit{et al.}\textsuperscript{125} proposed a similar binding mode for TSCs in the context of a CoMFA/CoMSIA analysis of the influence of ligand properties on the TSC-CZ interaction. The authors developed models with a high internal prediction capacity considering the contributions of the steric and electrostatic fields to the biological activity (54\% and 46\%, respectively) in the case of the CoMFA method. Moreover, the favourable steric maps have good complementarity with the S1 and S2 subsite properties on the CZ molecular surface. Taking compound 38 as a reference, the steric contour surrounding the ethyl substituent indicate that larger groups in these regions would improve the inhibitory potency. The electrostatic fields of CoMFA indicate a favourable interaction between the electronegative sulphur atom of the TSC moiety with the N atom of the Gln19 side chain. Considering that TSC derivatives are described as reversible covalent inhibitors of CZ,\textsuperscript{119,125} the proposed inhibition mechanism suggests (Figure 13.B) a nucleophilic attack of the thiolate anion of the Cys25 residue on the C atom of the TSC moiety assisted by the transfer of the His162 proton to the TSC sulphur. According to this model, the amino group of the side chain of Gln19 is close to the S atom of the TSC moiety to ensure a reasonable orientation of the H-bond and a possible role in the reverse reaction. The H-bond interaction could orient the S atom of the tetrahedral adduct in a favourable position to facilitate proton abstraction by residue His162, followed by electron transfer from the adduct to the thiolate anion of residue Cys25. More recently, kinetic studies allowed to infer the reversible nature of the inhibition of CZ by TSCs,\textsuperscript{126,127} establishing the classification of these ligands as “reversible covalent inhibitors”. The probable bioisosteric relationship between peptide bonds and TSC structural motifs represents an element that supports the proposed mechanism of action of TSCs.
In the case of the CoMSIA modelling, the results indicated a clear relationship between the steric, electrostatic, and hydrophobic descriptors and the biological activity (each accounting for 26%, 25% and 48%, respectively). The analysis of the steric contour maps surrounding the ethyl substituent in 38 leads to the same conclusions as obtained by CoMFA modelling. The contours near the imine nitrogen atom suggest that the presence of electropositive substituents at this position of the TSC group could improve the inhibitory potency.

Starting from TSCs derived from aryl-alkyl ketones and aryl-aldehydes, the analysis of the effect of structural modifications on the aromatic ring or the alkyl chain allowed to identify structural elements presents in TSCs with influence on activity as CZ/CZP inhibitors (SAR), as shown in Figure 14. Although the definition of pharmacophore is in progress, Jasinski et al. have developed a QSAR model that account the influence of electronic and steric properties of substituents in aromatic and lateral chain molecular regions, respectively, on the CZ inhibitory activity.

![Figure 14. Structural features present in aldehydes and alkyl aryl-ketones TSC derivatives as inhibitors of CZ.](image)

The main characteristics pointing toward a structure-activity relationship model for TSCs are as follows:
a) The systematic replacement of the ethyl group by smaller substituents at the "lateral region" in propiophenone derivatives conduct to a progressive reduction in activity along the TSC series (Et > Me > H). In parallel, the replacement by more voluminous substituents (for example, \( n \)-butyl) generally allows high activity values to be maintained.

b) Substitution of the ethyl group with a phenyl group gives the corresponding benzophenone TSC derivatives, some with IC\(_{50}\) values close to 0.020 \( \mu \)M. In parallel, the intercalation of spacers (e.g. -OCH\(_2\), -CH\(_2\)CH\(_2\)- or -CH=CH-) between one of the aromatic rings and the imine-type carbon of the TSC, allows the reduction of the molecular symmetry and planarity, giving rise to compounds that adopt a T-shaped conformation with increased affinity to CZ binding site.\(^{119,123b}\)

c) The incorporation of halogen atoms as substituents (Br, Cl), NO\(_2\) or CF\(_3\) in position 3 of the aromatic ring leads to high-potency compounds (except in the case of the incorporation of F), probably through the electronic effect on the TSC group reactivity and/or the affinity between the aromatic ring and S2 subsite. In this context, it is noteworthy that the simultaneous introduction of Cl atoms in positions 3 and 4 of the ring, or CF\(_3\) groups in positions 3 and 5, leads to very active compounds.

d) Introduction of phenyl, methoxy, phenoxy, amino or phenylamino groups as substituents in positions 3 or 4 in the aromatic ring leads to compounds with very low or no activity.
e) Derivatives incorporating substituents at position 2 of the aromatic ring (e.g., Br or Cl) result in compounds with low activity, regardless of the nature of the substitution at other ring positions. In this context, the cyclic derivatives where the alkyl chain at the "lateral region" is attached to position 2 of the aromatic ring could represent an exception to the rule. In this case, it is possible to observe compounds with IC$_{50}$ values close to 0.020 μM (TSC derivatives of 7-bromo-1-tetralone).$^{129}$

In 2011, Caputto $et$ $al.$$^{124}$ described the synthesis of 24 TSCs derived from 1-indanone as cyclic analogues of propiophenone TSC derivatives. They were anti-$T$. $cruzii$ with excellent selectivity indices (SI $>$ 150), indicating that these TSCs can be considered potential trypanocides. Some of them showed inhibitory activity against CZP (e.g. 5,6-dimethoxyindan-1-one N-(4-chlorophenyl) thiosemicarbazone with 67% of inhibition at 100 μM concentration), being the potential binding modes explored by molecular docking experiments, but no correlation was observed between the enzyme inhibition and its trypanocidal activity.

Several QSAR studies have been conducted to find correlations between TSC structural properties (represented by molecular descriptors) and their potency as CZ/CZP inhibitors. In this context, diverse strategies for QSAR model development have been applied, ranging from multiple-linear regression (MLR) models based on classical physicochemical descriptors,$^{128}$ CoMFA/CoMSIA$^{125}$ and HQSAR studies$^{130}$ to machine learning-based models based on techniques such as kNN that include topological descriptors.$^{131}$ These QSAR models served as a basis for development of new inhibitors, as illustrated by Jasinski $et$ $al.$ in the design of new TSCs with IC$_{50}$ values on the low nanomolar order.$^{128}$
Considering the proposed mechanism for the inhibitory activity of TSCs on CZ/CZP, it is clear that the TSC group plays a central role in the activity of these compounds. Therefore, the introduction of modifications to its structure can cause profound changes in the behaviour of the TSC derivatives as enzyme inhibitors. In this context, these modifications can be classified according to whether they include or not the formation of a heterocyclic ring.

Among the modifications that do not involve cyclization, the following can be highlighted:

a) Replacement of the S atom by potential bioisosteric groups (O, NH)

b) Reduction of the imino group

c) Derivatization of the -NH₂ group to give rise to secondary or tertiary thioamides.

d) Formation of metal complexes

In general, these modifications lead to compounds with low or no inhibitory activity. The replacements ‘S x O’ (semicarbazones)¹¹⁹ or ‘S x NH’ (aminoguanidines)¹³² lead to a very significant reduction in activity (especially in the ‘S x NH’ change). This sheds light on the importance of the elements that make up the TSC group considering the potential non-covalent intermolecular interactions and the reactivity of this group towards the Cys25 residue in CZ (see Figure 14).

Studies on the effect of coordination complexes of TSCs with gold, palladium, and platinum revealed that their trypanocidal action on *T. cruzi* is not mediated by CZP inhibition. The works of Carneiro et al.¹³³–¹³⁵ reported that whilst the Au (III) complexes were deleterious on both the trypomastigote and the amastigote forms of *T. cruzi*, these complexes showed a very low affinity for CZ. Furthermore, molecular docking simulation suggested that the Au(III) complexes weakly interact with CZ. Thus, and in parallel to what was pointed out in the case of Pd(II) and Pt(II) complexes with
trypanocidal effects,\textsuperscript{136} it is possible that the action of this class of compounds is mainly due to intracellular redox processes and formation of free radical species.

The following sections review heterocyclic derivatives evaluated as CZ/CZP inhibitors.

\textit{c) Heterocyclic compounds}

\textit{c1) Pentagonal heterocycles}

Conformational restriction, through ring formation, is a strategy frequently used in medicinal chemistry to study SAR and optimization of lead compounds.\textsuperscript{136} In this way, numerous authors\textsuperscript{119,123,137–139} have applied pharmacomodulation strategies to explore the impact of conformational constraints on the activity profile of TSC derivatives. through the formation of heterocyclic derivatives (see Figure 15). Therefore, it is possible to point out the following principal strategies:

a) Attachment of the alkyl chain at the “lateral region” to the -C(5)N(1)N(2)- substructure in the TSC group, giving rise to pyrazoline-type rings which retain the -C(=S)NH\textsubscript{2} group as a substituent on the ring structure. In this case, TSCs are considered classical acyclic bioisosteres of pyrazolines (43).

b) Inclusion of the substructure -C(3)S(6)N(4)- of the TSC group in thiazolidone-type rings, giving rise to 4-aryl/4-oxo thiazolylhydrazones and related compounds. In this case, TSCs are considered non-classical acyclic bioisosteres of these heterocyclic analogues (44, 45).
Figure 15. Heterocyclic compounds developed as cyclic bioisosteres of TSCs. The atoms of the TSC group and their counterpart in the resulting heterocycles are indicated in red.

The transformation of TSCs into pyrazolines (43), described by McKerrow et al, has led to compounds active on CZ. In general terms, the inhibitory potency of these compounds and the corresponding precursors (TSCs) are of the same order, providing trypanocidal effects and low levels of cytotoxicity in many cases.

The formation of thiazolylhydrazone derivatives (44, 45) implies the loss of the thiocarbonyl group, considered the target for a possible nucleophilic attack by Cys25. Therefore, it is expected that the mode of action of TSCs and thiazolylhydrazones show divergences. Starting from TSCs, Lima Leite et al. obtained 4-oxo thiazolylhydrazone derivatives that were active on *T. cruzi* epimastigotes and trypomastigotes. The authors attributed this activity to the inhibition of cysteine proteases based on molecular docking studies, making inferences about a probable difference in affinity and potency between stereoisomers (although without empirical corroboration since the
compounds were evaluated in vitro as racemic mixtures). Some years later, a small library of thiazolylhydrazones was tested both on CZ and T. cruzi parasites.\textsuperscript{137} In this case, the series included both 4-oxo and 4-thioxo thiazolylhydrazones to test the effects of an (isosteric) replacement of the carbonyl group at position 4 of the heterocyclic ring by a thiocarbonyl group (as a new potential site for a nucleophilic attack by Cys25) (Figure 15). Interestingly, one of the 4-thioxo analogues was significantly more potent than the corresponding 4-oxo precursor. Subsequent studies have suggested that the trypanocidal activity of 4-oxo thiazolylhydrazone-5-acetic acid derivatives involves CZ/CZP-independent mechanisms.\textsuperscript{141}

In 2014, Cardoso et al.\textsuperscript{142} published the evaluation of the anti-T. cruzi activity of 2-(pyridin-2-yl)-1,3-thiazoles derived from the corresponding TSCs, finding that many of these derivatives were CZ inhibitors, e.g. compound 45 (Figure 15), with a higher level of anti-T. cruzi activity than the TSCs from which they were derived. The development of thiazolylhydrazone analogues has made it possible to observe that the action profile of this class of compounds is similar to that of the 2-imino thiazolidin-4-ones, although frequently with low inhibitory potency on CZ.\textsuperscript{124,143–145}

Some other pentagonal heterocyclic compounds were also studied as inhibitors of CZP. It is interesting to highlight the derivatives that include in their structure an N-acylhydrazone (NAH) substructure and a heterocycle. Oxadiazole rings have been frequently selected as a building block considering a) their classification as a bioisostere of esters and amides, being more chemically stable than peptidic bonds, and b) their presence in previously observed antiparasitic compounds. In this context, dos Santos Filho et al.\textsuperscript{146} explored the NAHs as privileged structures included in a library of 16 compounds bearing the 3-(4-substituted-aryl)-1,2,4-oxadiazole scaffold. Among these, compounds 46 and 47a (Figure 16) demonstrated to be potent antitypansomal agents with low
toxicity both *in vitro* and *in vivo*. These derivatives were considered lead compounds in molecular docking studies on CZ, but experimental assays against the enzyme were not performed. Some years later, the same authors presented the structural design, synthesis, and anti-*T. cruzi* evaluation of new NAH-oxadiazole derivatives 47a-h, 48a-h and 49a-h, designed from a previous model of computational docking of oxadiazoles on CZ (serie 47 in Figure 16). The ability of these compounds to inhibit catalytic activity of the enzyme was tested, but there was no correlation between the enzyme inhibition and the antiparasitic activity of the compounds.
Figure 16. A. NAH-1,2,4-oxadiazole derivatives designed as CZ inhibitors with anti-*T. cruzi* activity. B. Strategy in the design of 1,3,4-oxadiazoles as CZ or *T. cruzi* inhibitors from esters and amide isosteres. In the case of compound 55, the pharmacophoric elements defined by Ref. 153 for this structural family have been indicated (hydrogen bond donors, hydrogen bond acceptors and aromatic/lipophilic groups in green, blue, and grey, respectively).

In 2009, Renslo *et al.*\(^{148}\) identified some reversible, non-covalent inhibitors of CZ through a virtual screening strategy based on molecular docking experiments. The optimization of 50 by replacement of the ester group by the 1,3,4-oxadiazole ring as a bioisosteric equivalent (Figure 16.B) gave rise to a series of 1,3,4-oxadiazoles with potencies improved by a factor of 500-fold. Detailed
investigation of the SAR series subsequently revealed that many members of the 1,3,4-oxadiazole class (and surprisingly also 50) act via divergent modes of inhibition (competitive or via colloidal aggregation) depending on the assay conditions employed. More recently de Souza et al.,149 based on the previous work, described the development of 2D QSAR and a 3D-QSAR-based pharmacophore from a series of inhibitors enriched in 1,3,4-oxadiazole derivatives (with 55 as a lead compound, IC$_{50}$ = 0.200 μM, Figure 16.B). As a result of the application of the mentioned computational methods in the context of molecular docking experiments, the study suggests the relevance of hydrogen bonding and π-π intermolecular interactions between the heterocyclic rings and sub-sites on the CZ surface, as well a possible way to improve the binding mode of these oxadiazole analogues by the introduction of bulky substituent that could fit into the S1 sub-site in CZ.

In 2013, ligand- and structure-based virtual screening methods were combined to explore the ZINC database in search of new CZ inhibitors.150 As a result, 12 compounds were identified as competitive non-covalent inhibitors of CZ in in vitro assays. The SAR analysis of these compounds, taking the analogue Neq42 (56) as a reference, identified the 2-acetamidothiophene-3-carboxamide substructure as fundamental for the inhibitory action on CZ and the trypanocidal effects for this series. Of note, the binding mode of one of the most potent analogues (e.g. Neq176 (57), considered a molecular simplification of Neq42, Figure 17) was established by X-ray crystallography (PDB ID: 4KLB). In 2015, Hoeltz et al.151 reported a 100 ns molecular dynamics study of Neq176 describing the mechanism of CZ inhibition in terms of hydrogen bonding interactions with the key residues Gly66, Met68, Asn69, and Leu160 and subdomain movements that close the active state of the enzyme.
In 2019, Andricopulo et al.\textsuperscript{152} described the molecular modelling, the synthesis, and the biological evaluation of cyclic imides as CZ inhibitors. Starting with a micromolar-range CZ inhibitor (58 at Figure 18, IC\textsubscript{50} = 2.2\,\mu M), a molecular optimization strategy resulted in the nanomolar-range inhibitor 59 (Figure 18, IC\textsubscript{50} = 0.6\,\mu M), which is highly active against \textit{T. cruzi} intracellular amastigotes (IC\textsubscript{50} = 1.0\,\mu M). The authors proposed a SAR scheme for the synthesised imide derivatives. Compound 58 structure was divided into five fragments, from which the most relevant SARs were as follows: a) the imide function is not essential for activity against \textit{T. cruzi}, b) the replacement of the 3-chloro-4-methoxyphenyl fragment is unfavourable for the activity on CZ and \textit{T. cruzi}, c) the removal of the ester group is unfavourable for activity against CZ, d) the replacement of the secondary amide proved to be tolerable regarding the trypanocidal activity, and e) the presence of other hydrophobic groups in place of the isopropyl group at compound 58 is essential for activity (e.g. benzyl group at compound 59, Figure 18).

\begin{figure}[ht]
\centering
\includegraphics[width=\textwidth]{molecular_structure.png}
\caption{Molecular structure of compounds Neq42 and Neq176.}
\end{figure}
Compounds containing pentagonal heterocyclic rings such as imidazole and benzimidazole are a remarkable group of CZ inhibitors. Some examples are reviewed below. In 2013, Porcal et al.\textsuperscript{153} reported the solid-phase synthesis of 33 1,2,5-tri-substituted benzimidazole derivatives and their \textit{in vitro} activity on CZP and \textit{T. cruzi} epimastigotes. Seven of these compounds were potent inhibitors of \textit{T. cruzi} growth with IC\textsubscript{50} values in the range 6–16 µM. Molecular docking studies revealed the binding orientation of the ligands in the active site of the enzyme providing new guidelines for the further design of better inhibitors. On the other hand, high throughput screening type processes on a library of almost 200,000 compounds resulted in the discovery of a benzimidazole type ligand, which demonstrated the ability to inhibit CZ in the nM order. In this case, the structure of the formed complex with one of these derivatives (compound 60 in Figure 19) was clarified by X-ray diffraction experiments.\textsuperscript{154} Subsequently, the use of strategies for the design of new analogues of the aforementioned compound allowed the identification of a series of ligands with greater potency on CZ. The study of these compounds in \textit{T. cruzi} allowed establishing that they are capable of generating trypanocidal effects at µM concentrations (Figure 19).\textsuperscript{155}

\textbf{Figure 18.} Cyclic imides as potential CZ/CZP inhibitors.
Figure 19. Imidazoles and benzimidazoles derivatives from Ref. 154-156.

In 2019, Fereira et al. reported a SAR study of benzimidazole CZ inhibitors by simulations and free energy calculations.\(^\text{162}\) MD simulations and free-energy calculations were used to shed light on qualitative SAR trends. Simulations revealed the most stable enzyme-ligand interactions and provided insights about enzyme selectivity. On the basis of the good overall agreement between calculated and experimental binding free energies, the current data provide a basis for employing similar calculations in prospective studies to guide potency optimization, in an effort to generate leads for Chagas disease treatment. In that work, the molecular requirements for the activity (Figure 20) and possible interactions between the compounds with the enzyme were proposed.

Figure 20. SAR of compound 60 derivatives according to reference 157.
Starting from the crystallographic structure of CZ forming a complex with benzimidazole 60, and a database made up of a total of almost 4 million compounds, a set of 18 structures was selected, whose activity against CZ was evaluated in in vitro experiments. Through these experiments it was possible to identify an imidazole compound capable of inhibiting the enzyme at concentrations of the order μM, although with marked cytotoxicity in mammalian cells (compound 62, Figure 20). The systematic modification of this ligand allowed to obtain more potent (nM order) and less cytotoxic compounds.

More recently, Medeiros et al. described a series of imidazoles that act as competitive and potent CZ inhibitors, using a combination of ligand- and structure-based drug design strategies. A series of 37 CZ inhibitors related to compound 62 were used to develop 2D and 3D predictive QSAR models such as HQSAR, AutoQSAR, CoMFA, and CoMSIA. The best QSAR models were examined together with the molecular docking results, leading to the proposal that the imidazole core of the compounds studied interacts with Trp184, and this was essential for the activity of this family of derivatives, pointing out the importance of establishing polar contacts at the solvent-exposed interface at S1/S1’ subsites. Considering that the studied compounds have two ring (one of them an imidazole ring) connected by a linker ideally having an extension of five atoms, including two H-bond acceptors and one H-bond donor group. This linker allowed the two aromatic rings to position correctly and interact optimally with the S1’ and S2 subsites on the CZ surface. These studies, also, highlighted the essential role that bulky groups play in the occupancy of the S2 subsite.

Similar findings were described by the same authors through a multiparameter optimization approach, molecular modelling, and SARs employed for the identification of some new
benzimidazole derivatives as potent competitive inhibitors of CZ with trypanocidal activity and suitable pharmacokinetics.

**c2) Hexagonal heterocycles**

In 2011,\textsuperscript{160} it was reported that some molecules having a purine or triazine core are potent non-peptide inhibitors of CZ. To gain insight into the structural requirements that may lead to enhanced activity of these molecules, CoMFA and CoMSIA studies of a series of purine-carbonitriles as CZ inhibitors were carried out. Semiempirical quantum calculations were used as a method to obtain reliable conformations for molecular alignment of the inhibitors within the CZ. Two different molecular alignments were used, resulting in 3 CoMFA models and 31 CoMSIA models. These models correspond to all possible combinations between five fields: steric, electrostatic, hydrophobic, hydrogen bond donor, and hydrogen bond acceptor. The contour maps obtained from these models show a preference toward the purine ring and indicate that bulky groups with a negative potential at the 3- and 5-positions of the phenyl ring are important structural requirements for inhibitory activity against CZ.

Pyrimidine and quinoline derivatives have shown action against different forms of the parasite and/or possible interactions with CZ. For example, it was described a pyrimidine \textsuperscript{161} and a quinoline derivative \textsuperscript{162} with IC\textsubscript{50} values of 9.1 μM and 68 μM against *T. cruzi*, respectively. \textbf{While possible interactions of pyrimidine derivatives with CZ were proposed on the basis of different computational studies, the inhibitory activity of some quinoline derivatives on CZ has been demonstrated experimentally but not seems correlated with the respective potency on *T. cruzi* parasites.}
In 2019, Fabian et al.\textsuperscript{163} evaluated the binding mode of ten quinoxaline compounds (64-67 in Figure 21) to a site adjacent to S2 (AS2) of CZ, according to Durrant et al. proposal.\textsuperscript{164} They were evaluated by a protocol that included a first analysis through docking experiments followed by a second analysis by MM-PBSA method. Through them it was demonstrated that quinoxaline compounds bearing substituents of different sizes at positions 3 or 4 of the heterocyclic ring might interact with the AS2. These compounds showed docking scores (\(\Delta G_{\text{dock}}\)) which were similar to those estimated for inhibitors that bind to the enzyme through non-covalent interactions. Nevertheless, the free binding energies (\(\Delta G\)) values estimated indicated that the derivatives 65 a-c (Figure 21), which bear bulky substituent at position 3 of the heterocyclic ring, became detached from the binding site under a dynamic study. Surprisingly, the evaluation of the inhibitory activity of CZP of some derivatives surprisingly showed that they increase the enzymatic activity. These results lead to the conclusion about the relevance of AS2 as a pocket for compound binding site, but not necessarily for the design of anti-chagasic compounds.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{quinoxalines.png}
\caption{Quinoxalines used in the docking study and in the analysis by MM-PBSA.}
\end{figure}

More recently, Ferreira et al.\textsuperscript{165} described the synthesis and biological evaluation of 22 analogs of \(N^4\)-benzyl-\(N^2\)-phenylquinazoline-2,4-diamine, previously described as competitive CZ inhibitor (\(K_i = 1.4 \ \mu\text{M}\)).\textsuperscript{170} These compounds contain modifications in the quinazoline core, and in
the substituents in positions 2 and 4 of this ring. These derivatives demonstrated low μM inhibition of the target proteases and trypanocidal activity against *T. cruzi* with low toxicity against myoblasts. During the optimization of the lead compound, structure-based design and prediction of physicochemical properties were employed to maintain potency against the enzyme. The global SAR, according to reference 169, for quinazoline compounds against CZ is shown in Figure 22.

![Figure 22](image)

*Figure 22.* Global SAR observed for *N*¹-benzyl-*N*²-phenylquinazoline-2,4-diamine derivatives, adapted from Ref. 165 and 166.

### 3. Other relevant inhibitors

The search for new CZ/CZP inhibitors has been expanded to the identification of compounds already approved as drugs for the treatment of other diseases, and which result appropriate for the treatment of Chagas disease. This approach, called “drug repositioning,” is attractive as it allows a significant reduction in time and cost in drug development.¹⁶⁷ The latter is a consequence of the use of compounds whose safety profile has already been considered acceptable in terms of regulatory requirements. In this context, some of the most important results obtained through the application of this strategy have been achieved by Bellera *et al.*,¹⁶⁸,¹⁶⁹ who have identified the drugs amiodarone, bromocriptine, clofazimine, benidipine and saquinavir as CZP inhibitors with a dose-dependent
trypanocidal effect (Figure 23). In this way, it has been shown that saquinavir can achieve more than 90% inhibition of the enzymatic activity of CZP in *in vitro* experiments, being active on *T. cruzi* practically regardless of the stage of its development.

The search for CZ/CZP inhibitors also includes the evaluation of natural compounds, present in plants and marine organisms. In this way, the identification of compounds of plant origin such as the sesquiterpenes α-copaene and zingiberene,170 and alkaloids such as cryptolepine171 (capable of inhibiting CZ activity at concentrations of the order μM), is of great interest for the development of analogues with a better pharmacodynamic profile (Figure 23).
Figure 23. A. Candidate drugs for repositioning for the treatment of Chagas disease through CZP inhibition. B. Compounds of natural origin with activity on CZ.

Regarding cryptolepine, in 2015,\textsuperscript{176} QSAR studies have been performed on 22 alkyldiamine cryptolepine derivatives, which could, at least in part, explain their antitrypanosomal activity. A multiple linear regression procedure was used to envisage the relationships between molecular descriptors and the CZ inhibitory activity of these derivatives. Results show high correlation between experimental and predicted activity values, indicating the validation and the good quality of the derived QSAR models. The developed QSAR models show that hydrophilic derivatives of cryptolepine have a good antitrypanosomal activity against CZ.
In the evaluation of products of natural origin, the discovery of compounds of marine origin capable of inhibiting CZ in concentrations of the order nM has been surprising. This is the case of compounds present in the coral *Plexaura homomalla* and the peptide metabolite gallinamide-A found in the cyanobacterium *Schizothrix* sp. This is not only capable of inhibiting CZ with IC$_{50}$ = 0.00026 µM but is also trypanocidal on the amastigote form of *T. cruzi*, with a value of IC$_{50}$ = 0.015 µM.

More recently, Proudst *et al.* starting from gallinamide-A and 23 synthetic analogues, evaluated against intracellular *T. cruzi* amastigotes and CZ, revealed that the N-terminal end of gallinamide-A weakly contributes in drug–target interactions. At the C-terminus, the intramolecular π–π stacking interactions between the aromatic substituents at P1′ and P1 restrict the bioactive conformation of the inhibitors, thus minimising the entropic loss associated with target binding. MD simulations showed that in the absence of an aromatic group at P1, the substituent at P1′ interacts with Trp184. The P1–P1′ interactions had no effect on anti-CZ activity, whereas anti-*T. cruzi* potency increased by fivefold, likely due to an increase in solubility/permeability of the analogues. Figure 24 shows the SAR of gallinamide-A and its derivatives, proposed by McKerrow in Ref. 165.
Figure 24. SAR of gallinamide-A and its derivatives as CZ/CZP inhibitors and trypanocide compounds, adapted from reference 165.

Recently, sulfonamides derived from anacardic acid and some chalcone derivatives\textsuperscript{176-179} have shown deleterious action against different stages of the parasite's life cycle and/or possible interactions with CZ through computational studies, but the inhibitory action of CZ/CZP has not yet been experimentally demonstrated.
Conclusions

Although Chagas disease is typical of underdeveloped countries, it constitutes a growing health problem in developed countries, given the migratory flows. In this context, the information collected in this review is valuable as a starting point to deepen the search for new inhibitors CZ/CZP, a validated target for the development of useful drugs for the treatment of Chagas disease. Many compounds of natural, synthetic, or semi-synthetic origin have demonstrated inhibitory action on CZ/CZP as a key enzyme in the development and survival of *T. cruzi*. Over the last 30 years, numerous peptide derivatives have demonstrated reversible or irreversible mechanisms of action on this cysteine protease and some of them have progresses to advanced stages of pre-clinical trials. However, non-peptide compounds containing functional groups considered bioisosteres of the peptide backbone (such as NAHs, TSCs and some heterocyclic rings) have been designed, synthesised, and evaluated as potential CZ/CZP inhibitors. Although many of them have demonstrated promising activity profiles on CZ/CZP at *in vitro* tests, all of them need to be further optimised in terms of selectivity and pharmacokinetic properties. It should be noted that there is no single feature for optimizing all CZ/CZP inhibitors. Each family requires different modifications to get closer to be serious drug candidates for clinical trials.

From this review emerged the following perspectives about the research on CZP inhibitors as potential drugs for treatments of Chagas disease:

a) the need to define early, in the research process, the relationship between the inhibitory activity on CZP/CZ and the trypanocidal profile of the compounds under investigation,

b) the evaluation of the inhibitors of CZ must be complemented with *in vitro* assays on polymorphic CZP extracted from parasites belonging to different DTUs,
c) the need to develop sufficiently selective inhibitors of CZ/CZP with minimal effects on mammalian cysteine proteases,

d) the benefit of the development of multi-target inhibitors active on more than one *T. cruzi* targets simultaneously, including CZP.

**Acknowledgments**

This work was financially supported by Agencia Nacional de Promoción Científica y Tecnológica, Universidad de Buenos Aires and Consejo Nacional de Investigaciones Científicas y Técnicas (Argentina) grants. MFM and AGM are members of the Research Career from National Scientific and Technical Research Council - Argentina (CONICET), Argentina.
Author Contributions

Gabriel Jasinski has a PhD in Pharmacy from the University of Buenos Aires (Buenos Aires, 2022). He is currently a Teacher in Medicinal Chemistry at the Faculty of Pharmacy and Biochemistry of the University of Buenos Aires, developing drug discovery studies against Chagas disease and cancer. His main research interests are related to computational and physical chemistry, applied to the design and study of mechanisms of action, as well as the synthesis and pharmacological evaluation of different classes of organic compounds.

María Florencia Martini is PhD in Biophysics and Biochemistry from the University of Buenos Aires (UBA, 2006). She is an Associate Professor at the Technological Institute of Buenos Aires (ITBA), developing projects applied to molecular biosurfaces and biological chemistry, with emphasis on methods and techniques of computational chemistry, in order to increase knowledge of the molecular behaviour of systems. She leads the Simulations and Molecular Modeling centre of the National Council for Scientific and Technical Research of Argentina (CONICET) at ITBA.

Albertina Gladys Moglioni is PhD in Organic Chemistry from the University of Buenos Aires (Buenos Aires, 1996). She is currently a Professor in Medicinal Chemistry at the Faculty of Pharmacy and Biochemistry of the University of Buenos Aires, developing projects applied to design and synthesis of compounds with possible pharmacological applications. She coordinates the Institute in Chemistry and Drug Metabolism (UBA-CONICET) and leads a research group related to lead optimization process, involving the design, synthesis, and pharmacological evaluation of different classes of organic compounds.
References

1. Fierz-David, H. E.; Beziehungen zwischen der chemischen Konstitution und der therapeutischen Wirksamkeit einiger künstlicher organischer Farbstoffe; Springer Berlin Heidelberg, 1935


7. Romaña, C.; Anales del Instituto de Medicina Regional (Tucumán) 1953, 3, 255.

8. Mazza, S.; Bassu, G.; Bassu, R.; Misión de Estudios de Patología Regional Argentina 1945


32. Ribeiro-Gomes, F. L.; Lopes, M. F.; DosReis, G. A.; *Negative signaling and modulation of macrophage function in Trypanosoma cruzi infection*; Landes Bioscience: Austin, Texas USA, **2007**


49. Zhai, X.; Meek, T. D.; *Biochemistry* 2018, 57, 3176. [https://doi.org/10.1021/acs.biochem.7b01250]


51. Arafet, K.; Świderek, K.; Moliner, V.; *ACS Omega* 2018, 3, 18613. [https://doi.org/10.1021/acsomega.8b03010]


76. Arafet, K.; Ferrer, S.; Moliner, V.; *Biochemistry* **2015**, *54*, 3381. [https://doi.org/10.1021/bi501551g]


82. DNDi; https://dndi.org/research-development/portfolio/k777/  


96. Rostovtsev, V. V; Green, L. G.; Fokin, V. V; Sharpless, K. B.; *Angewandte Chemie* 2002, 114, 2708. [https://doi.org/10.1002/1521-3757(20020715)114:14<2708::AID-ANGE2708>3.0.CO;2-0]


109. De Clercq, E.; Viruses 2010, 2, 1322. [https://doi.org/10.3390/v2061322]


126. Lill, I.; Trawick, M. L.; Investigation of Cruzain Inhibition by Thiosemicarbazone Compounds. Thesis. 2018


131. Rosas-Jimenez, J. G.; Garcia-Revilla, M. A.; Madariaga-Mazon, A.; Martinez-Mayorga, K.; *ACS Omega* 2021, *6*, 6722. [https://doi.org/10.1021/acsomega.0c05645]


177. Magalhães, E. P.; Gomes, N. D. B.; Freitas, T. A. de; Silva, B. P.; Ribeiro, L. R.; Ameida-Neto, F. W. Q.; Marinho,
M. M.; Lima-Neto, P. de; Marinho, E. S.; Santos, H. S. dos; Teixeira, A. M. R.; Sampaio, T. L.; Menezes, R. R. P. B.


179. de Brito, D. H. A.; Almeida-Neto, F. W. Q.; Ribeiro, L. R.; Magalhães, E. P.; de Menezes, R. R. P. B.; Sampaio,
T. L.; Martins, A. M. C.; Bandeira, P. N.; Marinho, M. M.; Marinho, E. S.; Barreto, A. C. H.; de Lima-Neto, P.; Saraiva,
[https://doi.org/10.1016/j.molstruc.2021.132197]