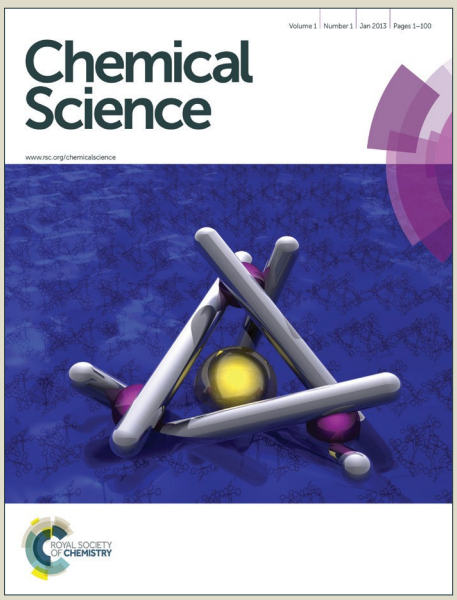
Chemical Science

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemicalscience

Chemical Science

ARTICLE

Iminoboronates are Efficient Intermediates for Selective, Rapid and Reversible *N*-Terminal Cysteine Functionalisation

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

DOI: 10.1039/x0xx00000x

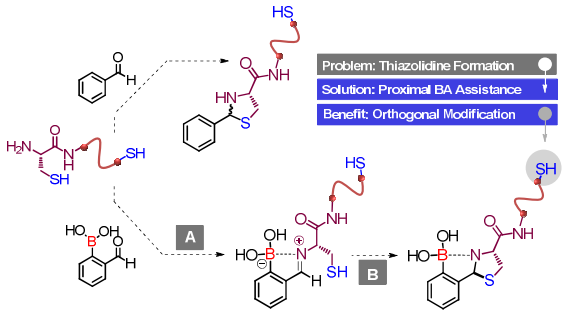
www.rsc.org/

Hélio Faustino,[a] Maria J. S. A. Silva,[a] Luís F. Veiros,[b] Gonçalo J. L. Bernardes,[c,d] Pedro M. P. Gois*[a]

We show that formyl benzeno boronic acids (2FBBA) selectively react with *N*-terminal cysteines to yield a boronated-thiazolidine featuring a B-N bond. The reaction exhibits a very rapid constant rate (2.38 ± 0.23 x 10^2 $M^{-1}s^{-1}$) under mild aqueous conditions (pH 7.4, 23 °C) and tolerates different amino acids at the position adjacent to the *N*-cysteine. DFT calculations highlighted the diastereoselective nature of this ligation reaction and support the involvement of the proximal boronic acid in the activation of the imine functionality and the stabilisation of the boronated-thiazolidine through a chelate effect. The 2FBBA reagent allowed the effective functionalisation of model peptides (C-ovalbumin and laminin fragment) and the boronated-thiazolidine construct was shown to be stable overtime though reversible in the presence of benzyl hydroxylamine. The reaction proved to be highly chemoselective, and 2FBBA was used to functionalize the *N*-terminal cysteine of calcitonin in the presence of a potentially competing in-chain thiol group. This exquisite selectivity profile enabled the dual functionalisation of calcitonin and the interactive orthogonal modification of this peptide when 2FBBA was combined with conventional maleimide chemistry. These results highlight the potential of this methodology to construct complex and well-defined bioconjugates.

Introduction

In recent years, the chemical functionalisation of peptides and proteins targeting the side chains of natural occurring amino acids developed into a powerful strategy to study fundamental biological processes, to construct therapeutic drugs and functional hybrid materials without the need for more specialized techniques.[1] Among the 20 canonic amino acids, cysteine (Cys) is a low abundant (<2%) residue that exhibits a highly reactive sulfhydryl side chain.[1,2] For these reasons, native or engineered Cys on the protein surface emerged as a preferred "hot-spot" for the site-selective modification of proteins and the construction of well-defined bioconjugates.[1-3] The functionalisation of this residue has been effectively achieved by promoting the thiol group elimination to generate dehydroalanine, *via* thiol-ene reaction or by exploring the thiol nucleophilicity in the presence of electrophiles.[1-3] These chemistries specifically target the sulfhydryl side chain and are mostly unselective when similar reactive thiol groups are present at the biomolecule's surface.[2] Therefore, the construction of more complex and well-defined bioconjugates, without resorting to mutant proteins,[4] the genetic encoding of specific amino acid sequences[5] or using multivalent reactive handles,[3b-d] is conditioned by the discovery of new chemical methods that may orthogonally functionalize Cys residues.[6]



Scheme 1. *N*-Terminal cysteine modification *via* thiazolidine and proposed proximal boronic acid assisted thiazolidine formation.

One of the most attractive ways to differentiate sulfhydryl side chains is to chemically target the *N*-terminal Cys residue with thioesters (Native Chemical Ligation),[7] aromatic cyanides[8] or aldehydes,[9] as these reagents selectively modify the 1,2-aminothiol in the presence of competing in-chain or C-terminal sulfhydryl side chains. Among these, the thiazolidine formation is a potentially powerful strategy to modify this residue,[9] because it uses as condensation partner, widely available, stable and structurally diverse aldehydes. Taking this into consideration, and with the exception of few examples,[9,10] thiazolidine formation reactions have not emerged as useful site-selective bioconjugation tools because the conditions

[a] *Research Institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, Universidade de Lisboa, Lisbon, Portugal. Email: pedrogois@ff.ulisboa.pt*
[b] *Centro de Química Estrutural, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal.*
[c] *Department of Chemistry, University of Cambridge, Lensfield Road, CN2 1EW, Cambridge.*
[d] *Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Avenida Professor Egas Moniz, 1649-028, Lisboa, Portugal.*
† Footnotes relating to the title and/or authors should appear here.

required for the condensation are typically incompatible with a wide range of biomolecules. Namely, this reaction proceeds at acidic pH 4-5, requires long reaction times (up to several days), multi-equivalents of reactants and if successful generates the thiazolidine as a mixture of diastereoisomers (Scheme 1).[1,9,10]

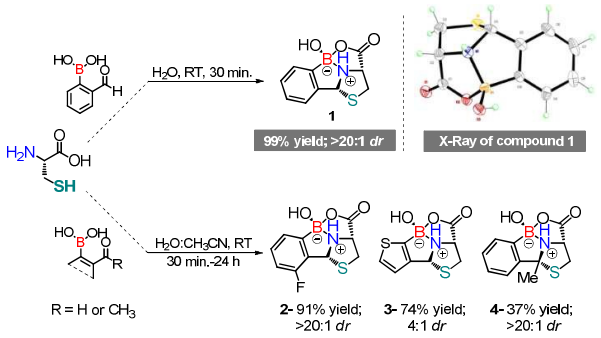
Results and discussion

Recognizing the potential of the thiazolidine method for orthogonal Cys bioconjugation, we initiated a study to improve this reaction profile. Recently we showed that 2-acetyl benzeno boronic acids (2ABBA) reversible functionalize protein exposed lysine residues *via* the formation of iminoboronates.[11] The boronic acid *ortho* to the carbonyl is key for the success of this bioconjugation as it promotes the imine formation and stabilizes the linkage by forming an N-B dative bond.[11] This proximal boronic acid assistance was recently used by others to enhance the formation of oximes, hydrazones and benzothiazoles.[12] Based on this, we reasoned that the condensation reaction of aldehydes with *N*-terminal Cys residues could also be significantly improved by generating a transient iminoboronate (Scheme 1A) *en route* to the cyclisation (Scheme 1B).

While submitting this manuscript, Gao *et al.* reported the labelling of *N*-terminal cysteines with 2FBBA. Whereas the disclosed data is consistent with our results, herein we considerably expand the scope of this reaction, exploring the systems orthogonality and reversibility in the selective dual labelling of peptides. In addition, we present a detailed study based on DFT calculations that highlight the invaluable role of the proximal boronic acid in the reaction mechanism.[12d]

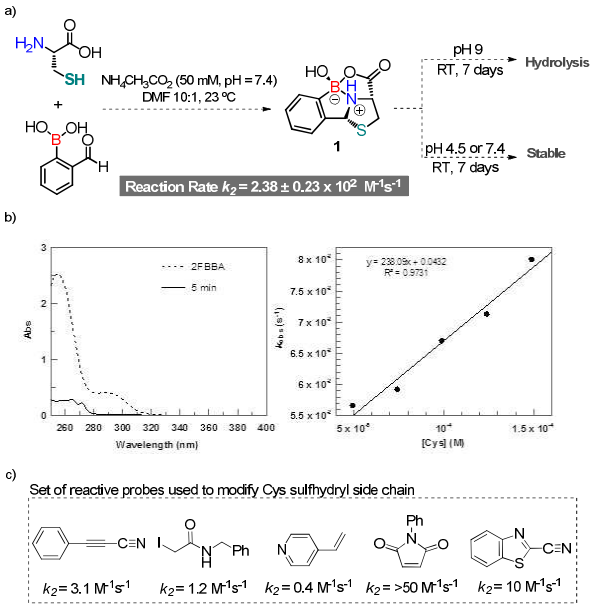
To test our idea, we started by performing a reaction between equimolar amounts of 2-formyl benzeno boronic acid (2FBBA) and Cys in water. To our surprise, the reaction proceeded smoothly at room temperature, and after 30 min. compound **1** precipitated from the reaction mixture. The desired product was simply collected by filtration with a 99% yield, high purity and almost as a single diastereoisomer (>20:1 dr). The unanticipated geometry of this tricyclic fused heterocycle, featuring a sp^3 boron centre enrolled in a N-B bond, was further elucidated based on the analysis of X-ray structure depicted on Scheme 2. Motivated by the efficiency of this assemblage, we tested the possibility of using different boronic acid scaffolds as shown in Scheme 2. Fluorination at the aromatic ring had no impact on the reaction (**2**) though, when using the 3-formyl-2-thienylboronic acid the condensation significantly lost its diastereoselectivity (**3**). Surprisingly, the 2ABBA also reacted under these rather mild conditions, yielding the heterocycle **4** with a fully substituted quaternary carbon centre, in 37% yield.

Next, we studied the impact of the proximal boronic acid on the condensation rate. To assess this, we monitored the reaction between of 2FBBA (200 µM) and Cys (1.5 equiv.) in acetate buffer (50 mM):DMF (10:1) at pH 7.4 using UV-Vis. As shown in Scheme 3A, this is a very fast reaction in which the aldehyde is readily consumed to yield the desired product in less than 5 min. In contrast, the reaction of benzaldehyde with L-Cys failed to deliver the thiazolidine under the same reaction conditions even when left for longer periods of time (see Figure S1 and S2, ESI).



Scheme 2. Reaction of Cys with different boronic acid scaffolds.

Interestingly, the free energy balances calculated for the two reactions by DFT,[13] corroborate those results with the reaction of 2FBBA with Cys being more favourable (by 17 kcal/mol) than the benzaldehyde one (see Figure S29, ESI). Product **1** is stabilized by the chelate effect on the boron atom, providing an extra thermodynamic drive to the reaction. These results demonstrated the importance of the neighbouring boronic acid for the success of the condensation reaction, hence the reaction rate was determined under pseudo first order conditions. As shown in Scheme 3b, the condensation reaction was confirmed as a very fast process exhibiting a rate constant of 2.38 ± 0.23 x 10^2 $M^{-1}s^{-1}$, which compares well with standard reagents used for bioconjugation of Cys residues (Scheme 3c). In particular, 2FBBA is 20 times faster than the cyanobenzothiazole (10 $M^{-1}s^{-1}$) that is commonly used for bioconjugation of terminal 1,2-aminothiols.[14]

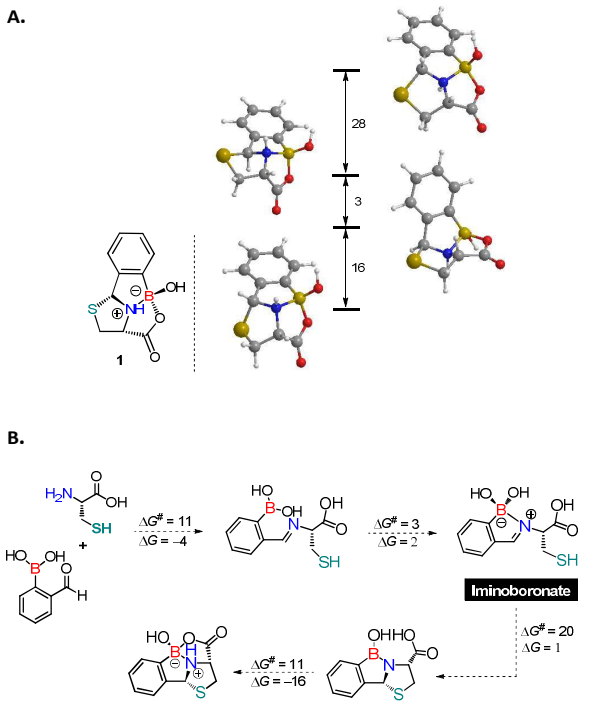


Scheme 3. a) Thiazolidine formation and stability at various pH; b) left - UV-vis absorption of 2-FBBA and reaction mixture at 5 min. right - k_{obs} values plotted against the concentration of cysteine to yield the second order rate

constant (k_2, $M^{-1}s^{-1}$) from the slope of the line. c) Probes for the modification of Cysteine.

Once we established the formation of this construct, we decided to also evaluate the heterocycle stability in ammonium acetate buffer. Despite exhibiting a rather complex structure, **1** proved to be stable (over 7 days) under acidic or neutral conditions, and only at pH 9, **1** slowly hydrolyzed to the individual components (see ESI, Figure S7).

As aforementioned, the thiazolidine condensation often yields conjugates as an inseparable mixture of diastereoisomers. Differently, the condensation between 2FBBA and Cys afforded the construct **1** in a highly diastereoselective manner. To elucidate this aspect of the reaction, the four more stable stereoisomers of product **1**, each with one different configuration on the N-atom or on the C2-atom of the thiazolidine ring were studied using DFT calculations.[15] The results obtained are depicted in Scheme 4A and indicate a clear preference for the diastereoisomer observed experimentally. This corresponds to the less strained structure, i.e., the one with a better balance of the ring tension due to the presence of three fused 5-member rings around the N-atom. The results in Scheme 4 show that the product obtained corresponds to the most stable diastereoisomer and indicate a thermodynamic control for the reaction.



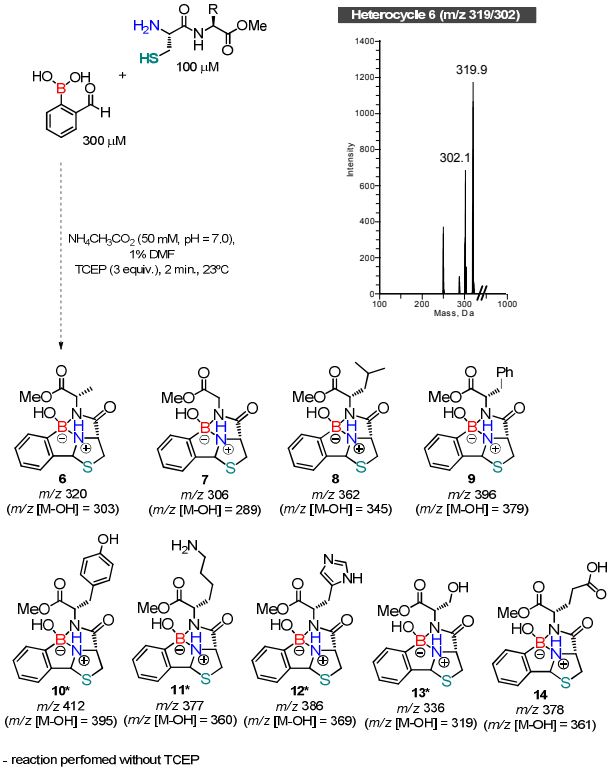
Scheme 4. A Stability difference between the four more stable diastereomers of product 1. **B** Schematic representation of the most important intermediates in the mechanism of formation of **1** from Cys and 2FBBA. Free energy values in kcal/mol.

Motivated by these results, we next used DFT calculations to elucidate the role of the proximal boronic acid in the reaction mechanism. A full account of the mechanism is given in the ESI including all calculated energy profiles (see ESI, Figures S30 and S31) and detailed representations of the path (see ESI, Schemes S1 and S2). The more important features of the proposed mechanism are presented in Scheme 4B. The mechanism calculated for the formation of **1** from 2FBBA and Cys comprises four main parts: the first corresponds to the condensation reaction between the amine group of *N*-Cys and the aldehyde of the boronic acid producing an imine involving the two initial reactants. This first section of the mechanism has a rather accessible barrier of $\Delta G^{\ddagger}$ = 11 kcal/mol and, overall, is exergonic by 4 kcal/mol, as expected for a reversible reaction. In the second part of the mechanism occurs the formation of a bond between the imine N-atom and the boron, resulting in a 5-member B-chelate ring. In this intermediate there is a tetravalent, formally negative, B-atom and an iminium group. This is a very facile step, almost barrierless ($\Delta G^{\ddagger}$ = 3 kcal/mol) and only slightly endergonic (ΔG = 2 kcal/mol). The activation of the C-atom of the iminium group, accomplished through the coordination of the nitrogen to the boron atom, makes possible the thiol attack and the resulting S–C in the next step of mechanism. Formation of the new S–C bond occurs simultaneously with proton transfer from the SH group to one of the OH groups attached to the B-atom and subsequent loss of the resulting water molecule. Thus, in the following intermediate there is a second 5-member ring fused to the first one by the C–N bond and a trivalent (neutral) B-atom stabilized by two coordinating atoms, the O-atom of the OH group and the N-atom of the amine. This section of the mechanism has the highest barrier of the entire path ($\Delta G^{\ddagger}$ = 20 kcal/mol) but it is essentially thermoneutral (ΔG = 1 kcal/mol). The last part of the mechanism corresponds to attack of the carboxylic group to the B-atom with formation of a new B–O bond and concurrent proton transfer to the amine N-atom. Thus, in the product **1** there is a third 5-member ring, fused with the previous two that is also another B-chelating ring. The barrier for this last section is moderate, $\Delta G^{\ddagger}$ = 11 kcal/mol, and the corresponding free energy balance is clearly exergonic (ΔG = –16 kcal/mol).

The overall path indicates a feasible reaction with a barrier of $\Delta G^{\ddagger}$ = 21 kcal/mol and clearly favourable free energy balance of ΔG = –19 kcal/mol, indicating the stability of the product as the main driving force for the reaction. Importantly, the B-atom has a two-fold role in the reaction mechanism. On one hand, it provides activation of the imine group by means of N–B coordination, promoting the formation of the S–C bond and, on the other, affords an additional stabilisation of the final product through multiple boron-coordination and the corresponding chelate effect.

Once we established the key features of this condensation, we evaluated the reaction between 2FBBA and model dipeptides featuring a *N*-terminal Cys. We started by reacting Cys-Ala-OMe (100 μM) with 3 equiv. of 2FBBA in ammonium acetate buffer (20 mM, pH 7.0) at 23 °C. Gratifyingly, the more complex structure of **5** was not detrimental for the reaction and the tricyclic heterocycle **6** (*m/z* 319/302) was readily assembled under these conditions in less than 5 min. Based on this result, other peptides were prepared and reacted with 2FBBA. As shown in Scheme 5, all peptides reacted efficiently under these conditions and the assemblage was complete

within 5 min. at 23 °C. Peptides constructed with amino acids such as glutamate (Glu), tyrosine (Tyr), serine (Ser) or Lys presenting functionalities that could in principle disturb the heterocycle framework assemblage, were relatively innocent in this process demonstrating the robustness and generality of this bioconjugation method (see ESI, figures S8-16).



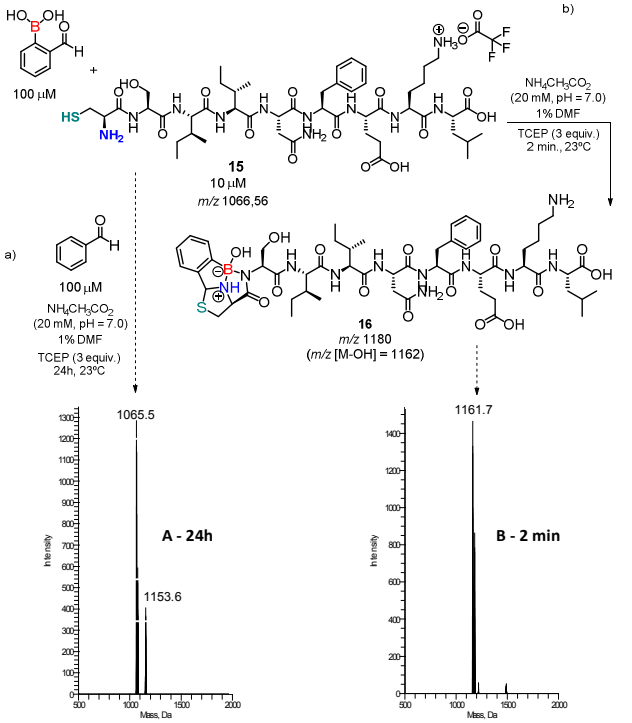
Scheme 5. Conjugation of 2FBBA with a variety of *N*-terminal Cysteine dipeptides.

Encouraged by these results, an ovalbumin derived peptide exhibiting an *N*-terminal Cys residue was constructed and submitted to conjugation with 10 equiv. of 2FBBA. As shown in Scheme 6, the expected construct (*m/z* 1180/1162) was immediately formed at pH 7 and remained stable in solution up to 24 h in this media (See ESI, Figure S17). Similar results were obtained when using only 3 equiv. of 2FBBA reagent (Scheme 6). Interestingly the presence of a Lys residue did not affect the assemblage of **16**, probably due to a higher reversibility of the iminoboronate formed with the lysine ε-amino group.

To further demonstrate the importance of the proximal boronic acid for the conjugation, benzaldehyde was reacted with peptide **15** and the reaction was monitored by ESI-MS. As shown in Scheme 6A only after 24 h the thiazolidine (*m/z* 1154) was detected in the reaction mixture, corroborating that the standard thiazolidine condensation is not a useful bioconjugation tool under these conditions.

Then we tested the condensation of 2FBBA with a laminin fragment that exhibits several residues that may interfere directly with the boronic acid function. As shown in scheme 7, despite the presence of a Glu adjacent to the *N*-terminal or in chain Tyr and Ser residues, the assemblage with 2FBBA proceeded smoothly to yield the expected construct (*m/z* 1081/1063) under the optimized reactions conditions.

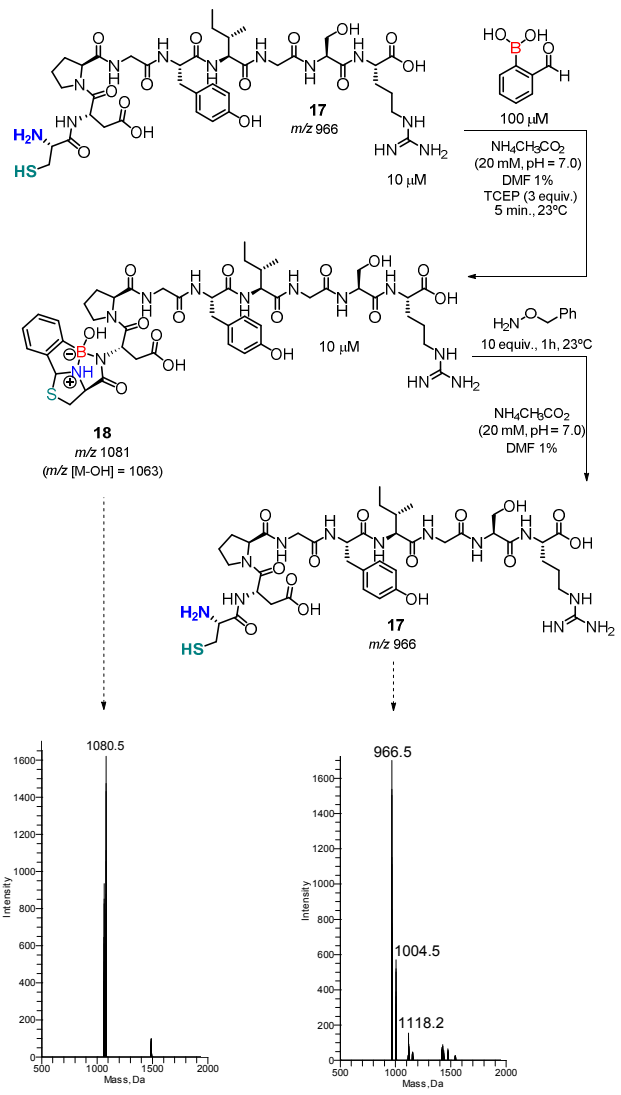
The thiazolidine is a known reversible linkage in the presence of hydroxylamine derivatives. Therefore, and to understand if this key feature of thiazolidines is retained in our newly formed constructs, we decided to treat compound **18** with 10 equiv. of benzyl hydroxylamine. As expected the construct was fully hydrolyzed to the corresponding parent peptide.



Scheme 6. a) Reaction of C-ovalbumin with benzaldehyde; b) Reaction of C-ovalbumin with 2FBBA.

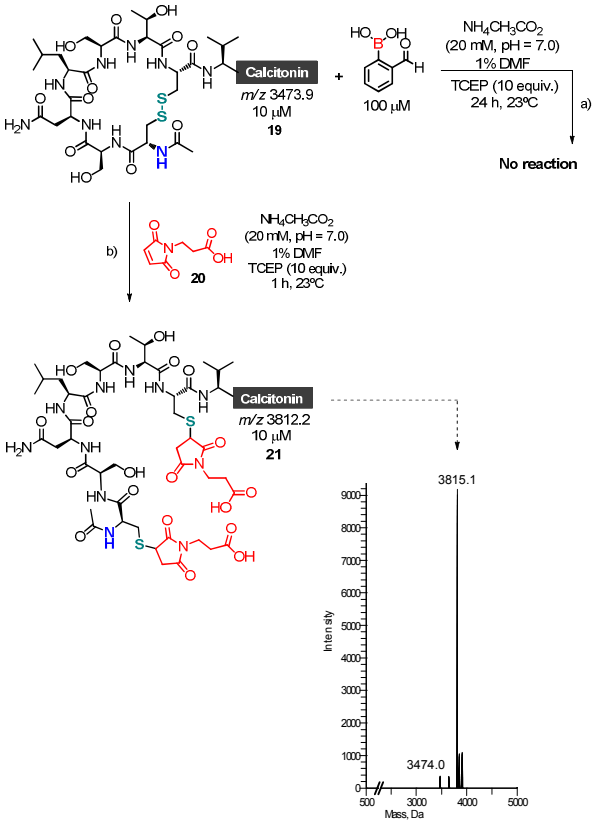
Once we established the assemblage of Cys *N*-terminal peptides with 2FBBA, the reaction was evaluated on peptides featuring in-chain Cys residues. To study this, the therapeutic peptide *N*-acetylated calcitonin was treated with TCEP to reduce the disulphide bridge and then first reacted with a model maleimide and subsequently with 2FBBA. As expected, the exposed thiol groups were efficiently alkylated with maleimide **20** (Scheme 8). In contrast, under the same reducing conditions, the peptide failed to react with the boronic acid scaffold suggesting that acetylation at the nitrogen precludes the generation of the construct with 2FBBA and supports the formation of the iminoboronate *en route* to the cyclisation. To further explore this observation, the non-acetylated calcitonin **22** was submitted to conjugation with 2FBBA in the presence of TCEP (Scheme 9). Gratifyingly, under these conditions, the reaction proceeded efficiently affording

the conjugate **23** in less than 5 min. Next, we assumed that the formation of **23** occurred without reaction of the in-chain thiol group, making it available for one additional round of functionalisation. Hence, conjugate **23** was treated with maleimide **20**, and this simple protocol yielded the conjugate **24**, orthogonally modified at both Cys residues (Scheme 9).



Scheme 7. Condensation of 2FBBA with a laminin fragment and its reversibility promoted by benzyl hydroxylamine.

Finally the possibility to revert the functionalisation operated with 2FBBA with benzyl hydroxylamine (Scheme 7) offers the possibility to use this scaffold as a protective group of the *N*-terminal Cys and engaged in an interactive orthogonal dual modification. To test this, the conjugate **23** was functionalized with a PEG-maleimide, and subsequently treated using 20 equiv. of benzyl hydroxylamine (Scheme 9). Then, the boron construct was effectively removed exposing the *N*-terminal sulfhydryl side chain that was simply modified now using maleimide **28** (Scheme 10).

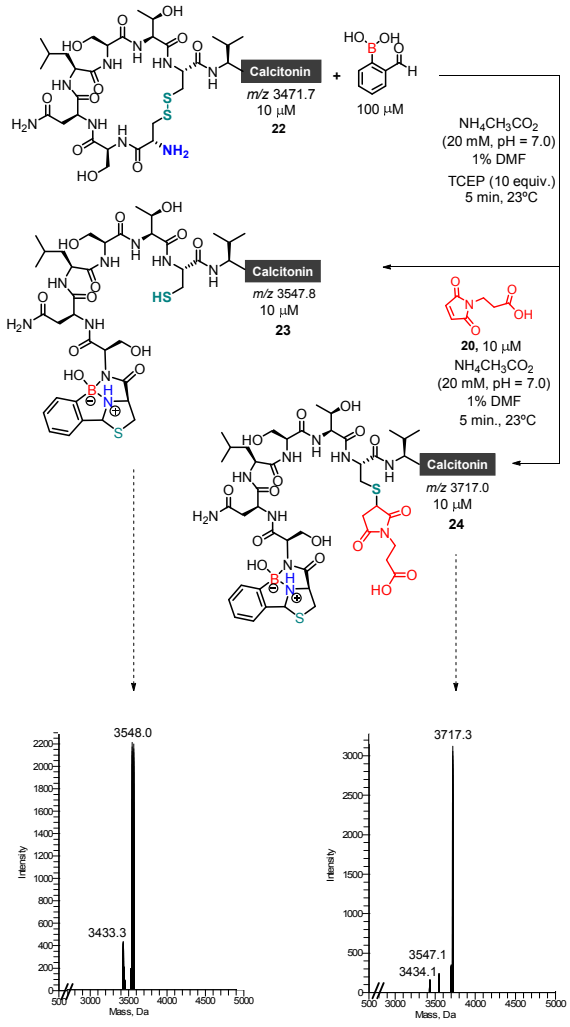


Scheme 8. a) acetyl salmon fail to give any product in the presence of 2FBBA; b) alkylation of the exposed thiol groups with maleimide **20**.

Conclusions

In this study we showed that 2FBBA effectively reacts with the 1,2-aminothiol function of *N*-terminal Cys to generate a boronated-thiazolidine featuring a B-N bond. The reaction is very rapid ($2.38 \pm 0.23 \times 10^2$ $M^{-1}s^{-1}$) and proceeds under mild conditions (pH 7.4, 23 °C) at a near stoichiometric ratio of reagents. DFT calculations performed on this system showed that the proximal B-atom has a dual role. On one hand, it provides activation of the imine group by means of N–B coordination, promoting the formation of the S–C bond and, on the other, affords an additional stabilisation of the final product through multiple boron-coordination and the corresponding chelate effect. Regarding the reaction scope, the functionalisation proceeds efficiently with structurally diverse *N*-terminal Cys dipeptides (e.g., Gly, Tyr, Ser, or Lys), bearing functionalities that could in principle disturb the heterocycle framework assemblage. Similarly, model peptides including C-ovalbumin and laminin fragment reacted effectively with 2FBBA. The boronated-thiazolidine constructs were shown to be stable in acidic conditions, neutral pH and slightly basic conditions (< pH 9) though fully reversible upon the addition of benzyl hydroxylamine. The reaction was shown to be highly selective and 2FBBA was used to functionalize the *N*-terminal Cys of calcitonin in the presence of an in-chain thiol

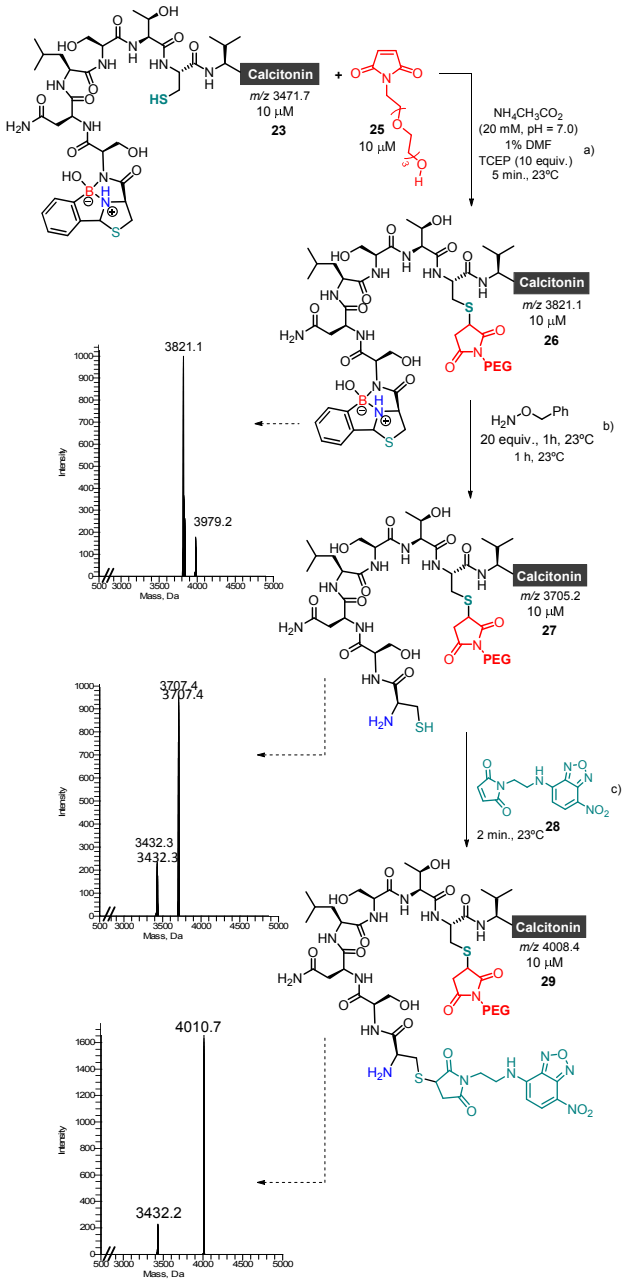
group. This selectivity profile was further explored in a dual functionalisation of calcitonin with a boronated-thiazolidine/PEG-maleimide and in an interactive installation of two different maleimides onto this peptide. The enclosed results showcase 2FBBA reagents as powerful tool to selectively functionalize biomolecules and unravel a new strategy for the orthogonal construction of more complex and well-defined bioconjugates.



Scheme 9. Selective modification of *N*-terminal and in chain cysteines with 2FBBA and maleimide **20** respectively.

Acknowledgements

We acknowledge the financial support of Fundação para a Ciência e a Tecnologia (FCT) Portugal (grants SFRH/BPD/102296/2014, PTDC/QEQMED/5512/2014, PTDC/QEQ-QOR/1434/2014; P.M.P. Gois and G. J. L. Bernardes are FCT Investigators; iMed.ULisboa grant UID/DTP/04138/2013); and the EU (Marie-Sklodowska Curie ITN *ProteinConjugates*. L.F.V. acknowledges Fundação para a Ciência e Tecnologia, UID/QUI/00100/2013. G.J.L.B. is a Royal Society University Research Fellow. The X-ray crystal structure determination reported in this paper was carried out by Dr Andrew D. Bond (Dept. of Chemistry, Univ. Cambridge).



Scheme 10. a) Selective modification of *N*-terminal and in chain Cys with 2FBBA and maleimide **24** respectively; b) removal of 2FBBA by addition of benzyl hydroxylamine; c) selective modification of *N*-terminal Cys with maleimide **28**.

Notes and references

1 Selected reviews on chemical protein modification: a) Y. Huang and L. Liu, *Nat. Chem.*, 2016, **8**, 101; N. Krall, F. P. da Cruz, O. Boutureira and G. J. L. Bernardes, *Nat. Chem.*, 2016, **8**, 103; b) D. A. Shannon and E. Weerapana, *Curr. Op. Chem. Bio.*, 2015, **24**, 18; c) O. Boutureira and G. J. L. Bernardes, *Chem. Rev.*, 2015, **115**, 2174; d) O. Konievab and A. Wagner, *Chem. Soc. Rev.*, 2015, **44**, 5495; e) Y. Takaoka, A. Ojida and I. Hamachi, *Angew. Chem. Int. Ed.*, 2013, **52**, 4088; f) N. Stephanopoulos and M. B. Francis, *Nat. Chem. Biol.*, 2011, **7**, 876.

2 a) J. M. Chalker, G. J. L. Bernardes, Y. A. Lin and B. G. Davis, *Chem. Asian J.*, 2009, **4**, 630; b) J. M. Chalker, S. B. Gunnoo, O. Boutureira, S. C. Gerstberger, M. F.-González, G. J. L. Bernardes, L. Griffin, H. Hailu, C. J. Schofielda and B. G. Davis, *Chem. Sci.*, 2011, **2**, 1666. Reports dealing with the selective labeling of different cysteines in a protein are rare and based on the differences of the environment where the cysteine is located, see: c) V. Ratner, E. Kahana, M. Eichler and E. Haas, *Bioconjug. Chem.*, 2002, **13**, 1163. d) P. Moody, V. Chudasama, R. I. Nathani, A. Maruani, S. Martin, M. E. B. Smith and S. Caddick, *Chem. Commun.,* 2014, **50**, 4898.

3 a) P. M. S. D. Cal, G. J. L. Bernardes, and P. M. P. Gois, *Angew. Chem. Int. Ed.*, 2014, **53**, 10585; b) M. Rashidian, S. C. Kumarapperuma, K. Gabrielse, A. Fegan, C. R. Wagner and M. D. Distefano, *J. Am. Chem. Soc.*, 2013, **135**, 16388. c) A. Maruani, M. E. B. Smith, E. Miranda, K. A. Chester, V. Chudasama and S. Caddick, *Nat. Commun.*, 2015, **6**, 6645. d) T. Wang, A. Riegger, M. Lamla, S. Wiese, P. Oeckl, M. Otto, Y. Wu, S. Fischer, H. Barth, S. L. Kuan and T. Weil, *Chem. Sci.*, 2016, **7**, 3234.

4 a) Q.-Ying Hu, F. Bertib and R. Adamo, Chem. Soc. Rev., 2016, **DOI:** 10.1039/C4CS00388H; b) D. M. Patterson, L. A. Nazarova, and J. A. Prescher, *ACS Chem. Biol.*, 2014, **9**, 592; c) J. T. Ngo and D. A. Tirrel, Acc. Chem. Res., 2011, **44**, 677; d) J. Xie and P. G. Schultz, *Nat. Rev. Mol. Cell Biol.*, 2006, **7**, 775.

5 C. Zhang, M. Welborn, T. Zhu, N. J. Yang, M. S. Santos, T. V. Voorhis and B. L. Pentelute, *Nat. Chem.*, 2016, **8**, 120.

6 Selected examples: a) A. Sachdeva, K. Wang, T. Elliott and J. W. Chin, *J. Am. Chem. Soc.*, 2014, **136**, 7785; b) D. Zhang, N. O. Devarie-Baez, Q. Li Jack, R. Lancaster Jr. and M. Xian, *Org. Lett.*, 2012, **14**, 3396; c) M. Mühlberg, M. G. Hoesl, C. Kuehne, J. Dernedde, N. Budisa and C. P. R. Hackenberger, Beilstein, *J. Org. Chem.* 2015, **11**, 784; d) J. Kim, M.-H. Seo, S. Lee, K. Cho, A. Yang, K. Woo, H.-S. Kim and H.-S. Park, *Anal. Chem.*, 2013, **85**, 1468; e) J. Grünewald, D. H. Jones, A. Brock, H.-P. Chiu, B. Bursulaya, K. Ng, T. Vo, P. Patterson, T. Uno, J. Hunt, G. Spraggon and B. H. Geierstanger, *ChemBioChem*, 2014, **15**, 1787; f) R. I. Nathani, P. Moody, V. Chudasama, M. E. B. Smith, R. J. Fitzmaurice and S. Caddick, *Chem. Sci.*, 2013, **4**, 3455.

7 P. Dawson, T. Muir, I. Clark-Lewis and S. Kent, *Science*, 1994, **266**, 776.

8 H. Ren, F. Xiao, K. Zhan, Y.-P. Kim, H. Xie, Z. Xia and J. Rao, Angew. Chem. Int. Ed., 2009, 48, 9658; Y. Yuan and G. Liang, *Org. Biomol. Chem.*, 2014, **12**, 865.

9 L. Zhang and J. P. Tam, *Anal. BioChem.*, 1996, **233**, 87;

10 a) G. Casi, N. H.-Dezot, K. Zuberbühler, J. Scheuermann, and D. Neri, *J. Am. Chem. Soc.*, 2012, **134**, 5887; b) G. J. L. Bernardes, M. Steiner, I. Hartmann, D. Neri and G. Casi, *Nat. Prot.*, 2013, **8**, 2079.

11 a) P. M. S. D. Cal, J. B. Vicente, E. Pires, A. V. Coelho, L. F. Veiros, C. Cordeiro and P. M. P. Gois, *J. Am. Chem. Soc.*, 2012, **134**, 10299; b) P. M. S. D. Cal, R. F. M. Frade, V. Chudasama, C. Cordeiro, S. Caddick and P. M. P. Gois; *Chem. Commun.*, 2014, **50**, 5261; c) P. M. S. D. Cal, R. F. M. Frade, C. Cordeiro and P. M. P. Gois; *Chem. Eur. J.*, 2015, **21**, 8182.

12 a) P. Schmidt, C. Stressa and D. Gillingham, *Chem. Sci.*, 2015, **6**, 3329; b) A. Bandyopadhyay and J. Gao, *Chem. Eur. J.* 2015, **21**, 14748; O. Dilek, Z. Lei, K. Mukherjee and S. Bane, *Chem. Commun.,* 2015, **51**, 16992; c) A. B. Draganov, K. Wang, J. Holmes, K. Damera, D. Wang, C. Dai and B. Wang, *Chem. Commun.*, 2015, **51**, 15180. d) A. Bandyopadhyay, S. Cambray and J. Gao, Chem. Sci., 2016. DOI: 10.1039/C6SC00172F.

13 a) R. G. Parr, W. Yang, Density Functional Theory of Atoms and Molecules, Oxford University Press, New York, 1989; b) DFT calculations at the M06-2X/[6-31+G(d,p), PCM] level were performed using the Gaussian 09 package. A complete account of the computational details and the corresponding list of references are provided as Supporting Information.

14 O. Koniev, G. Leriche, M. Nothisen, J.-S. Remy, J.-M. Strub, C. S.-Reiss, A. V. Dorsselaer, R. Baati and Alain Wagner, *Bioconjugate Chem.*, 2014, **25**, 202.

15 In the more stable isomers, the configuration of the boron centre is mostly dictated by the neighbour N- and by the aminoacid C_{α}.