

Organic & Biomolecular Chemistry

Accepted Manuscript

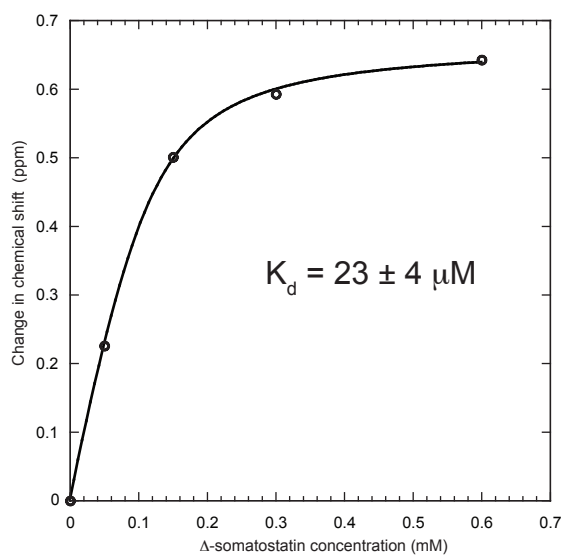
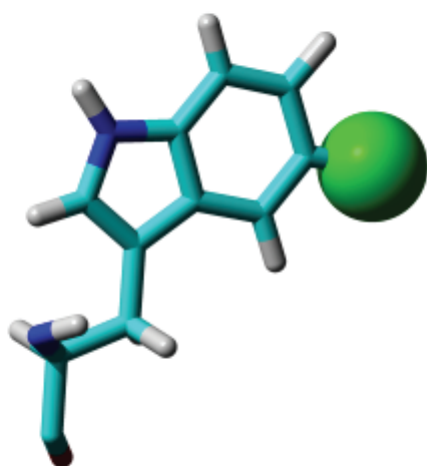


This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



Fluoroindole recombinant protein labelling enables a ^{19}F NMR study to observe protein-ligand binding and dissociation constant determination

COMMUNICATION

^{19}F NMR spectroscopy monitors ligand binding to recombinantly fluorine-labelled $b'x$ from human protein disulphide isomerase (hPDI)

Cite this: DOI: 10.1039/x0xx00000x

Received 00th January 2012,
Accepted 00th January 2012Rose Curtis-Marof,^a Denisa Doko,^a Michelle L. Rowe,^a Kirsty L. Richards,^a
Richard A. Williamson^{a,*} and Mark J. Howard^{a,*}

DOI: 10.1039/x0xx00000x

www.rsc.org/

We report a protein-observe ^{19}F NMR-based ligand titration binding study of human PDI $b'x$ with Δ -somatostatin that also emphasises the need to optimise recombinant protein fluorination when using 5- or 6-fluoroindole. This study highlights a recombinant preference for 5-fluoroindole over 6-fluoroindole; most likely due to the influence of fluorine atomic packing within the folded protein structure. Fluorination affords a single ^{19}F resonance probe to follow displacement of the protein x -linker as ligand is titrated and provides a dissociation constant of $23 \pm 4 \mu\text{M}$.

Protein disulphide isomerase (PDI) is a key enzyme responsible for the formation of native disulphide bonds in proteins that enter the secretory pathway of eukaryotic cells. PDI is a multifunctional protein able to catalyse the oxidation and isomerisation of disulphide bonds, as well as to bind to unfolded proteins and act as a molecular chaperone.¹ The isomerisation of incorrectly paired cysteine residues is often a rate-limiting step on the folding pathway of disulphide bond-containing proteins both *in vitro* and *in vivo*.² The ability of PDI to combine redox and molecular chaperone-like activities allows it to bind to partly structured folding intermediates and to catalyse simultaneously protein folding and associated native disulphide bond formation.³ PDI is the archetype of a large family of ER-resident PDI-like proteins.²

PDI contains four thioredoxin-like domains, two of which -- like thioredoxin itself -- have redox-active catalytic sites (**a** and **a'**) and two of which do not (**b** and **b'**). The domain order is **a-b-b'-x-a'-c**, where **x** is a 19 residue linker between the **b'** and **a'** domains^{4,5} and **c** is a C-terminal acidic tail containing the KDEL ER-retention signal. The **a** and **a'** domains are responsible for the redox activity of PDI while the **b'** domain has been shown to be essential for ligand binding.⁶ The **b'** domain binds both small and large peptide ligands, although large ligands also require the **a** and **a'** domains to contribute to the overall binding interaction.^{6,7} The **b'** domain has

been shown to contribute to the substrate specificity of PDI^{4,8} and is required for disulphide isomerisation reactions in protein substrates.⁹

The nature of substrate binding involving the **b'** domain continues to be a subject of interest and any information regarding the role of **b'** and **x** is extremely useful. Recent studies have highlighted the **b'** domain of human PDI (hPDI) to be structurally connected and influenced by **x**, the linker that connects the **b'** and **a'** domains. The **b'x** construct has been shown to be receptive to ligand binding⁴ and **x** has also been shown to moderate homodimerisation in **b'x** and **bb'x**.¹⁰ These data supported the previously published crystal structure studies of a **b'x** mutant that confirms **b'** contains the thioredoxin-fold with **x** occluding the ligand binding site as mapped using NMR chemical shifts.^{4,11} As a result, it is thought that **b'x** exists in two conformational states with the **x-linker** 'capping' or 'uncapping' the ligand binding site and that 'uncapping' provided opportunity for homodimerisation.^{4,10} Limited proteolysis, using wild-type and mutant proteins to promote or retard 'capping', confirmed the **x-linker** operates similarly in full-length hPDI.¹² Homodimerisation in hPDI is a mechanism by which the binding site is hidden from substrates, and has been proposed as a potential regulation mechanism for the protein *in vivo*.¹⁰

hPDI **bb'x** and **bb'** have been studied using ^1H , ^{15}N -HSQC chemical shift mapping to provide estimated dissociation constants (K_d) with peptide ligands Δ -somatostatin (AGSKNFFWKFTSS) and mastoparan (INLKALAALAKKIL)^{4,8} with these studies approximating the K_d in the range of 0.1 – 1.0 mM. Despite these investigations, the role of the **x-linker** remains a point of conjecture because backbone resonances in the linker itself were notoriously difficult to observe and track in the ^1H , ^{15}N -HSQC ligand binding studies mentioned above due to chemical shift line broadening. Therefore, we present this concise study of **x** within **b'x** using ^{19}F NMR to follow the displacement of the 'capped' form of the linker when a known peptide ligand is titrated against constant protein concentration. Switching to ^{19}F NMR

COMMUNICATION

provides an optimum detection window of the **x-linker** to observe both ligand binding and displacement of the linker from the binding site. This is facilitated by the fact that hPDI **b'x** contains only a single tryptophan residue that resides at position 347 within the **x-linker** and positioned directly above the reported ligand binding site as shown in Figure 1. In addition, we also describe the rationale behind our choice of indole fluorination that provides optimal expression.

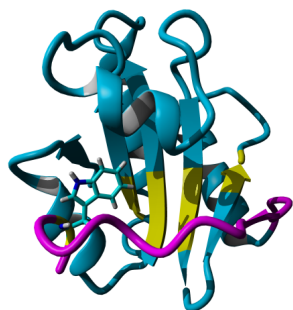


Figure 1. Crystal structure of hPDI **b'x** I272A with the 'capped' **x-linker** shown in magenta and the ligand-binding site in yellow. Trp 347 is shown in stick form. The sequence and secondary structure of **b'x** is shown in ESI Figure 1.

All NMR datasets were acquired at 298K using a Bruker Avance III 14.1 T (600 MHz ^1H) NMR spectrometer equipped with a 5mm QCI-F cryoprobe. Data were processed using Bruker Topspin 3.1 software and referenced using the position of the $^1\text{H}_2\text{O}$, relative gyromagnetic ratios (for ^{15}N) or using trifluoroacetic acid (for ^{19}F). Bacterial expression and purification of all proteins were identical to that reported previously for hPDI **b'x**^{4, 10} with additional steps only to permit fluorination in accordance with the recently published protocol by Crowley and co-workers.¹³ For convenience, the entire growth protocol and purification can be found in the ESI. The method utilises a standard bacteria minimal growth media with 60 mg/L of indole, 5-fluoroindole or 6-fluoroindole added 15 minutes prior to induction with IPTG (isopropyl β -D-1-thiogalactopyranoside). All indoles were added to the growth media from a 200x dimethylsulfoxide (DMSO) stock solutions. An SDS-PAGE gel illustrating the effect of each indole on expression of hPDI **b'x** is shown in Figure 2.

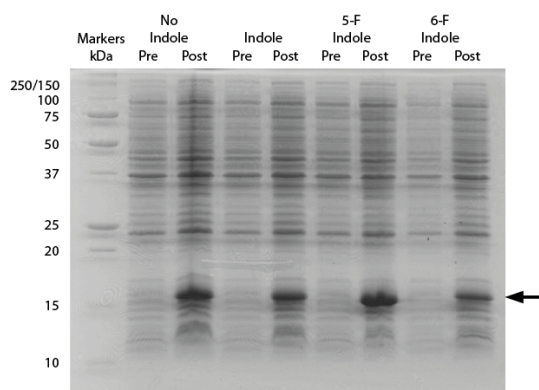


Figure 2. 15% SDS-PAGE of hPDI **b'x** showing samples taken pre- and post-induction in minimal media without indole, with indole, with 5-fluoroindole (5-F) and with 6-fluoroindole (6-F). Lane 1 shows Precision-Plus protein standard markers labelled with molecular weight (kDa) and the black arrow highlights the position of **b'x**.

The inclusion of indole using DMSO as a stock carrier liquid did not hinder induction of the protein and the addition of 5-fluoroindole did also not alter the protein yield. From a minimal growth medium both indole and 5-fluoroindole provided yields of 6-8 mg/L of purified protein. However, Figure 2 highlights the effect of using 6-fluoroindole where the yield of recombinant **b'x** was found to be consistently lower than for indole and 5-fluoroindole with an estimated yield of ca. 0.5 mg/L.

Differences in fluoroindole overexpression are likely be related to the positioning of the fluorine in the indole ring and the SDS-PAGE gel in Figure 2 highlights the importance of utilising fluorination that supports correct folding of a stable protein. This hypothesis was further corroborated using mass spectrometry (Bruker MicroTOF-Q) where 5-F-Trp incorporation was found to be >80% but 6-F-Trp incorporation was ca. 55% (see ESI Figure 2) confirming that 6-F-Trp incorporation was less successful. In an attempt to understand the effect of fluorination, Figure 3 uses the 'capped' **b'x** I272A structure with *in silico* modified tryptophans to highlight the structural effects of 5-fluoro- (5-F-Trp: Fig 3a) or 6-fluorotryptophan (6-F-Trp: Fig 3b) labelling of **b'x**.

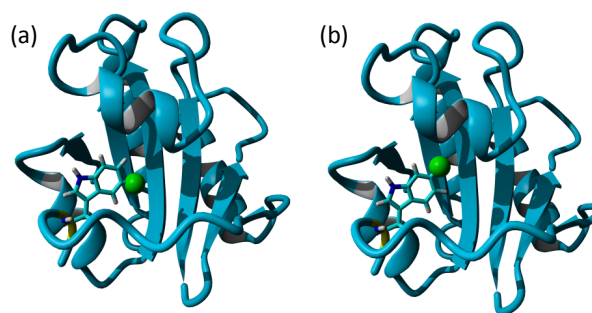


Figure 3. Structure of hPDI **b'x** I272A (3bj5.pdb) showing the **x-linker** tryptophan modified *in silico* with fluorine inserted into the indole ring as a green van der Waals sphere to represent 5-F labelling (a) and 6-F labelling (b).

Close inspection of these structures suggests that when fluorine is present as 6-F-Trp, the fluorine is packed against residues F223, I284, F287 and F288 (see ESI Figure 3a). These four residues were emphasised as providing key interactions between **b'** and the **x-linker** in the crystal structure.¹¹ In contrast, inspection of the structure where fluorine is present using 5-F-Trp confirms that this structure experiences fewer critical interactions because the fluorine atom resides in a pocket above the ligand-binding site (Figure 3a and ESI Figure 3b). The structural stability difference expected for 5-F-Trp versus 6-F-Trp is a likely explanation for the difference in expression yields when using 5-fluoro- and 6-fluoroindole. It could be suggested that 6-fluoroindole may have compromised cell metabolism resulting in the loss of production. However, observed growth rates were similar regardless of indole utilised and close inspection of post-expression gel bands in Figure 2 between 20-100 kDa supports comparable background protein levels across all samples. This observation would support that 6-fluoroindole does not adversely affect tryptophan synthase because many of these background proteins must also require tryptophan. In addition, we can confirm from other studies that 6-fluoroindole can be used to successfully

produce other proteins with 6-F-Trp (unpublished results). However, these data do highlight the need to carefully consider the choice of fluorination sites when producing labelled proteins. The structural integrity of wild-type **b'x** expressed in the presence and absence of 5-F-Trp was assessed using $^1\text{H},^{15}\text{N}$ -HSQC data obtained from samples grown in minimal media containing a single nitrogen source of 0.6 mg/L $\text{U-}^{15}\text{N}$ ammonium sulphate. $^1\text{H},^{15}\text{N}$ -HSQC **b'x** and **5-F-Trp-b'x** spectra are shown in Figure 4. The spectra overlay extremely well and suggest both samples have the same fold and structural arrangement, which is confirmed via a minimal chemical shift map analysis of the data (ESI-Figure 4). Furthermore, these data are also comparable to previously published $^1\text{H},^{15}\text{N}$ -HSQC spectra of wild-type and **x-linker** 'capping' mutants of **b'x**.^{4,11}

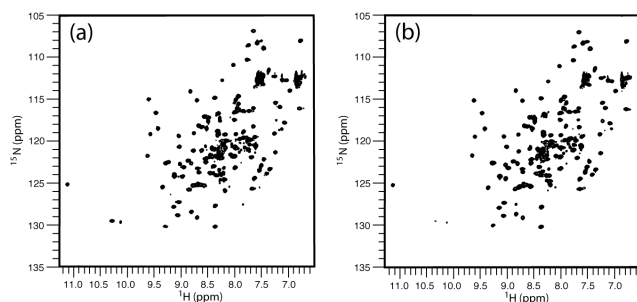


Figure 4. $^1\text{H},^{15}\text{N}$ -HSQC spectra of hPDI **b'x** (a) and **5-F-Trp-b'x** (b) from 0.25 mM protein samples. Data were acquired for each spectrum as 2048 x 256 points over 45 minutes.

The degree of fluorination was estimated by comparing the relative ratios of the indole NH peaks in the $^1\text{H},^{15}\text{N}$ -HSQC spectra of **b'x** and **5-F-Trp-b'x** and was found to vary between 80-90% depending on the preparation. This is in good agreement with the mass spectrometry data shown in ESI Figure 2. ^{19}F NMR spectra of 0.3 mM hPDI **5-F-Trp-b'x** were obtained with increasing Δ -somatostatin peptide concentration and are shown in Figure 5. Each 1D ^{19}F NMR experiment was acquired over 1,024 scans using a relaxation delay of 2 s and an acquisition time of 0.321 s. The data obtained in Figure 5 demonstrates the utility of the QCI-F cryoprobe and the ability to create a dataset within half the time required to acquire a $^1\text{H},^{15}\text{N}$ -HSQC dataset. This is especially important because **b'x** is prone to dimerise over time and the monomeric state in the predominantly 'capped' **x** conformation is required for this measurement. It is for this reason that a spectrum of a 48-hour 'aged' **b'x** sample is also shown in Figure 4 where both 'capped' and 'uncapped' forms can be distinguished by ^{19}F NMR. Using a fresh, purified sample enabled the titration of Δ -somatostatin to be monitored through a single ^{19}F resonance that reports on the **x-linker** status as the peak tracks from the 'capped' toward the 'uncapped' state. It is interesting to note the reduction in ^{19}F resonance line width as ligand is added; this supports the displacement of the **x-linker** in **b'x** as ligand binds to the protein. Once the linker is displaced, it will display increased mobility compared to the bulk protein, which manifests as an increase in the ^{19}F T_2 relaxation time and subsequent reduction in line width. This reduction in linewidth associated with **x-linker** displacement has also been observed in $^1\text{H},^{15}\text{N}$ HSQC spectra.^{4,11} The change in ^{19}F chemical shift with ligand concentration from Figure 5 can be plotted as shown in Figure 6 and fitted using the well-documented equation

below^{14,15} where Δ is the observed change in chemical shift, Δ_0 is the maximum shift and $[L]$ and $[P]$ are ligand and protein concentration respectively.

$$\Delta = \Delta_0 \frac{(K_d + [L] + [P]) - \sqrt{((K_d + [L] + [P])^2 - 4[P][L])}}{2[P]}$$

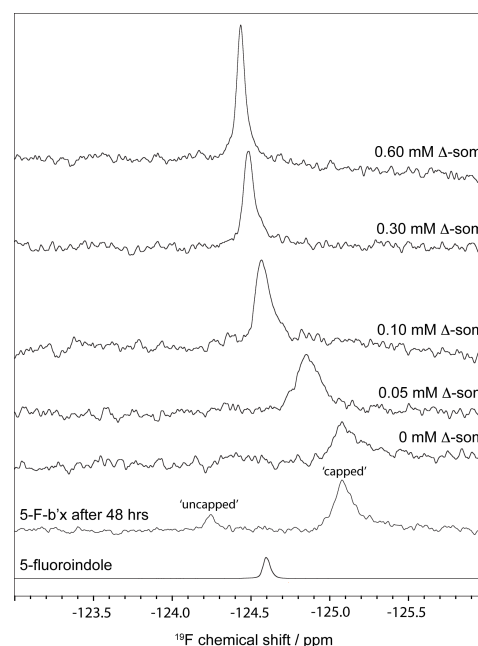


Figure 5. ^{19}F NMR spectra of 0.3 mM **5-F-Trp-b'x** with increasing concentration of Δ -somatostatin peptide (Δ -som). ^{19}F NMR spectra of 48 hr 'aged' **5-F-trp-b'x** shows ^{19}F resonances reporting on 'capped' and 'uncapped' **x-linker** and 5-fluoroindole are also shown.

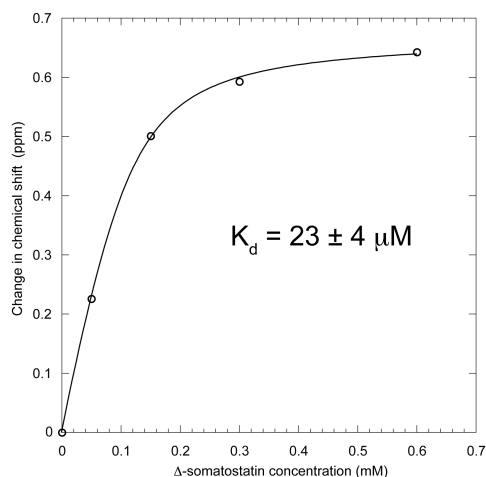


Figure 6. Titration saturation curve of 0.3 mM **5-F-Trp-b'x** when increasing the concentration of Δ -somatostatin peptide-ligand. Fitting the points to the equation provides the curve and solution for K_d of $23 \pm 4 \mu\text{M}$.

Solving the equation to fit the curve in Figure 6 produces a dissociation constant for hPDI **5-F-Trp-b'x** with Δ -somatostatin of $23 \pm 4 \mu\text{M}$. This value is marginally lower than that obtained from earlier studies that report K_d is in the 1.0-0.1 mM range and suggests the affinity of **5-F-Trp-b'x** for Δ -somatostatin is higher than for the wild-type **b'x** protein. This likely to be due to a small but significant destabilisation of 'capped' species caused by fluorination that

promotes displacement of the **x-linker** and exposure of the binding site. Fluorine substitutions in proteins and peptides have been known to stabilise and drive hydrophobic interactions and the ligand-binding process across many PDIs including hPDI **b'x** is considered as hydrophobically driven.

If fluorine mediated hydrophobicity increased the strength of binding of the **x-linker** region to **b'** then one might expect 'uncapping' to be more unfavourable with the consequence that ligand affinity is reduced. This does not appear to be the case.

A third explanation could be that the ligand experiences favourable hydrophobic interactions with the 'uncapped' **x-linker**; this would manifest as increased affinity and a lower K_d . These observations suggest the smaller K_d is driven by destabilisation of the 'capped' state and/or the 'uncapped' state experiences a hydrophobically driven binding event that is stabilised by fluorination. The latter hypothesis proposes a potential role for the **x-linker** in ligand binding and such a process has been suggested from analysis of recent crystal structures of yeast¹⁶ and hPDI¹⁷ as well as *in silico* modelling and SAXS analysis of PDI from *Humicola insolens*.^{18,19}

Conclusions

¹⁹F labelling using 5-fluoroindole to produce 5-fluorotryptophan for protein NMR spectroscopy is straightforward and successful. In addition, 5-fluoro-L-tryptophan is approximately 50-times more expensive than 5-fluoroindole per gram, thus making fluorindole-based labelling economical. However, our experience draws attention to the care required to find the optimum fluorindole for the system to be studied. hPDI **b'x** was best served using 5-fluoroindole but it is worth noting that 4-fluoro-, 5-fluoro- and 6-fluoroindoles are all available from general chemical suppliers. As a result, we recommend small-scale expression tests with different fluorindoles to discover the optimum labelling strategy. Our results also suggest that interrogation of a known structure can help identify the optimum fluorindole through analysis of close-quarter atomic interactions.

Ultimately, ¹⁹F NMR has provided a useful shift in the NMR timescales of detection that was a challenge regarding **x-linker** observation using ¹H-¹⁵N HSQCs. Fluorine NMR enables direct observation of the tryptophan side chain in **b'x** and the ¹⁹F chemical shift change can be used to study ligand binding and for identification of the 'capped' and 'uncapped' conformations of the protein. ¹⁹F NMR has also confirmed that the **x-linker** is displaced upon binding as the 5-F-Trp chemical shift tracks from the 'capped' to 'uncapped' state.

The data provides an excellent fit and a K_d of 23 μ M for Δ -somatostatin binding to **5-F-b'x** and this value is marginally lower than 0.1-1 mM reported for the wild-type protein. We suggest this was most likely due to destabilisation of 'capped' **5-F-b'x**, stabilisation of protein-peptide complex or a combination of both factors. Ultimately, this approach can be expanded further to study multiple PDI domain constructs and other PDI family members to monitor structure-function of this isomerase and folding chaperone.

Notes and references

^a Protein Science Group, School of Biosciences, University of Kent, Giles Lane, Canterbury, Kent, CT2 7NJ, UK. E-mail: r.a.williamson@kent.ac.uk or m.j.howard@kent.ac.uk

We would like to thank Mr. Kevin Howland for performing the mass spectrometry at Kent, The Wellcome Trust for supporting this project through Equipment Grant 091163/Z/10/Z (MJH and RAW) and Project Grant 093125/B/10/Z (MJH and RAW) and Prof. Robert Freedman and Dr. Katrine Wallis for discussions and comments regarding this manuscript.

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/c000000x/

1. F. Hatahet and L. W. Ruddock, *Febs J*, 2007, 274, 5223-5234.
2. F. Hatahet and L. W. Ruddock, *Antioxid Redox Signal*, 2009, 11, 2807-2850.
3. R. B. Freedman, P. Klappa and L. W. Ruddock, *Embo Reports*, 2002, 3, 136-140.
4. L. J. Byrne, A. Sidhu, A. K. Wallis, L. W. Ruddock, R. B. Freedman, M. J. Howard and R. A. Williamson, *Biochem J*, 2009, 432, 209-217.
5. A. Pirneskoski, P. Klappa, M. Lobell, R. A. Williamson, L. Byrne, H. I. Alanen, K. E. Salo, K. I. Kivirikko, R. B. Freedman and L. W. Ruddock, *J Biol Chem*, 2004, 279, 10374-10381.
6. P. Klappa, L. W. Ruddock, N. J. Darby and R. B. Freedman, *Embo Journal*, 1998, 17, 927-935.
7. P. Koivunen, K. E. Salo, J. Myllyharju and L. W. Ruddock, *J Biol Chem*, 2005, 280, 5227-5235.
8. A. Y. Denisov, P. Maattanen, C. Dabrowski, G. Kozlov, D. Y. Thomas and K. Gehring, *Febs J*, 2009, 276, 1440-1449.
9. N. J. Darby, E. Penka and R. Vincentelli, *J Mol Biol*, 1998, 276, 239-247.
10. A. K. Wallis, A. Sidhu, L. J. Byrne, M. J. Howard, L. W. Ruddock, R. A. Williamson and R. B. Freedman, *Protein Sci*, 2009, 18, 2569-2577.
11. V. D. Nguyen, K. Wallis, M. J. Howard, A. M. Haapalainen, K. E. Salo, M. J. Saaranen, A. Sidhu, R. K. Wierenga, R. B. Freedman, L. W. Ruddock and R. A. Williamson, *J Mol Biol*, 2008, 383, 1144-1155.
12. L. Wang, S. J. Li, A. Sidhu, L. Zhu, Y. Liang, R. B. Freedman and C. C. Wang, *J Biol Chem*, 2009, 284, 199-206.
13. P. B. Crowley, C. Kyne and W. B. Monteith, *Chem Commun (Camb)*, 2012, 48, 10681-10683.
14. M. P. Williamson, *Prog Nucl Magn Reson Spectrosc*, 2013, 73, 1-16.
15. L. Fielding, *Current topics in medicinal chemistry*, 2003, 3, 39-53.
16. G. Tian, S. Xiang, R. Noiva, W. J. Lennarz and H. Schindelin, *Cell*, 2006, 124, 61-73.
17. C. Wang, W. Li, J. Q. Ren, J. Q. Fang, H. M. Ke, W. M. Gong, W. Feng and C. C. Wang, *Antioxidants & Redox Signaling*, 2013, 19, 44-53.
18. O. Serve, Y. Kamiya, A. Maeno, M. Nakano, C. Murakami, H. Sasakawa, Y. Yamaguchi, T. Harada, E. Kurimoto, M. Yagi-Utsumi, T. Iguchi, K. Inaba, J. Kikuchi, O. Asami, T. Kajino, T. Oka, M. Nakasako and K. Kato, *J Mol Biol*, 2010, 396, 361-374.
19. M. Nakasako, A. Maeno, E. Kurimoto, T. Harada, Y. Yamaguchi, T. Oka, Y. Takayama, A. Iwata and K. Kato, *Biochemistry*, 2010, 49, 6953-6962.