

Synthesis of Next-Generation Maleimide Radical Labels

Bouchra Hajjaj^{a,b,1}

Anokhi Shah^{a,b}

Stacey Bell^{a,b,2}

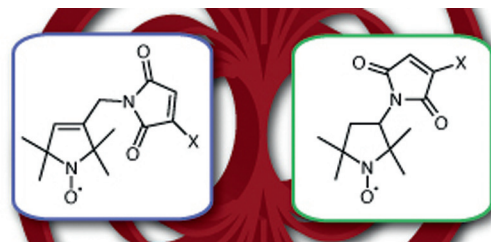
Sally L. Shirran^b

Catherine H. Botting^{b,c}

Alexandra M. Z. Slawin^c

Alison N. Hulme^d

Janet E. Lovett^{*a,b}



^a SUPA School of Physics and Astronomy, University of St Andrews, St Andrews, KY16 9SS, UK
jel20@st-andrews.ac.uk

^b BSRC, University of St Andrews, St Andrews, KY16 9ST, UK

^c EaStChem School of Chemistry, University of St Andrews, St Andrews, KY16 9ST, UK

^d EaStChem School of Chemistry, University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK

Received: 01.04.2016

Accepted after revision: 28.05.2016

Published online: 05.07.2016

DOI: 10.1055/s-0035-1562451; Art ID: st-2016-d0228-l

Abstract The synthesis and characterization of four new nitroxide-radical-containing next-generation maleimides are presented. Each new label has a single leaving group which is either a phenoxyl or bromide. The linker between the maleimide and the nitroxide-containing framework is either a racemic mixture of a short chain or an achiral longer chain. These molecules have been designed to site-specifically label vicinal cysteines in proteins for magnetic resonance studies. The characterization of the final products includes crystallography and the labeling of sperm whale myoglobin protein.

Key words next-generation maleimide, succinimide, nitroxide spin label, radical, EPR spectroscopy, PRE NMR, cysteine labeling, site-directed spin labeling

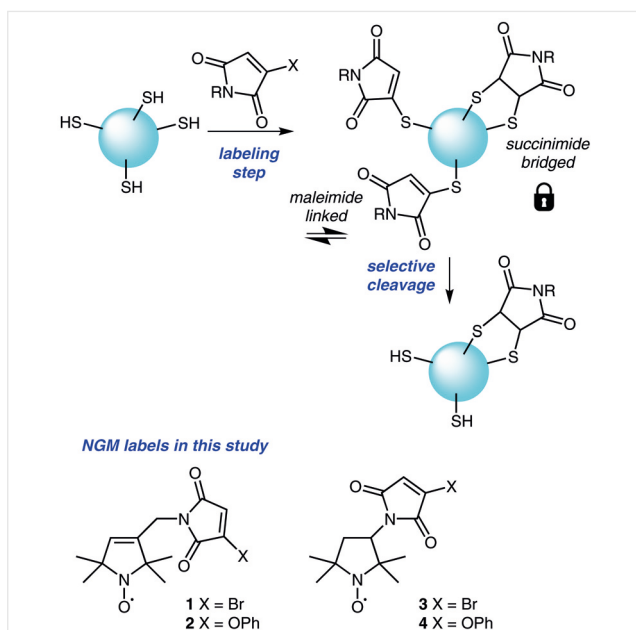
Spin labels are small molecules that attach to larger molecules of interest and that stabilize a free radical. The spin labels are most likely to be used to report on their microenvironment, accessibility, motional dynamics, or to enable nanometer-scale distance measurements by electron paramagnetic resonance (EPR) spectroscopy.³ Alternatively, they may be used as paramagnetic relaxation enhancing reagents in nuclear magnetic resonance (NMR) spectroscopy.⁴ It is therefore often desirable to have a spin label that demonstrates little conformational freedom while having enough flexibility to avoid disruption of the protein fold.

Commonly these labels stabilize a nitroxyl radical as part of a five- or six-membered carbon framework and are hence called nitroxide spin labels. These are functionalized to react at the site of interest. For proteins this site is often the thiol of cysteine. The labels can be made to covalently attach reversibly or irreversibly with short or long linkers.

Attachment chemistries utilize iodocetamides and maleimides to give carbon–sulfur bonds and thiosulfonates which result in disulfide-linked reagents.⁵

Proteins which are excreted from the biological cell rarely contain cysteines suitable for labeling. Therefore cysteines can be engineered into sites of interest in proteins and labeled using site-directed spin labeling (SDSL).³ However many excreted proteins contain disulfide bonds, where two cysteines are close in space rather than necessarily sequence, and intracellular proteins contain cysteines much more frequently.⁶ These situations are not ideal for designing SDSL experiments. One method to overcome this problem is to incorporate unnatural amino acids into the protein sequence. These have bioorthogonal reactivities and can therefore undergo selective labeling, alternatively there is a published example of a nitroxyl radical being incorporated directly.^{3,7} This field is progressing but at the moment the technology cannot be readily applied to any protein, and the spin labels are on long flexible linkers which is a disadvantage for many applications.

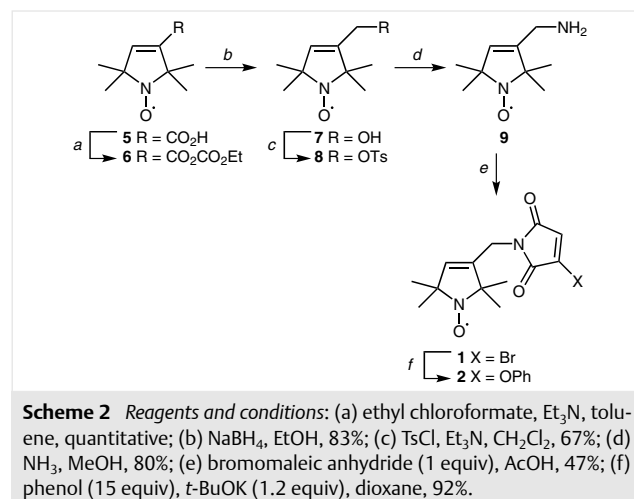
Here we report on the synthesis of spin labels based on next-generation maleimides.⁸ Following literature precedence, these should bind to pairs of cysteines which are close in space ('vicinal') via a succinimide bridge but bind to single cysteines with a sulfur-maleimide bond which can be reversed (Scheme 1). The labels proposed here are different to the existing next-generation maleimide spin labels, TPMP and TPM_cP proposed by Baker and co-workers as spinostic reagents.⁹ In that work they did not use succinimide bridging but instead had two leaving groups to result in a maleimide-bridged product. Here we wish to create spin labels that will give a different chemical reactivity, depending on whether single or pairs of cysteines are bound, to enable selective labeling of only pairs of cysteines (Scheme 1).



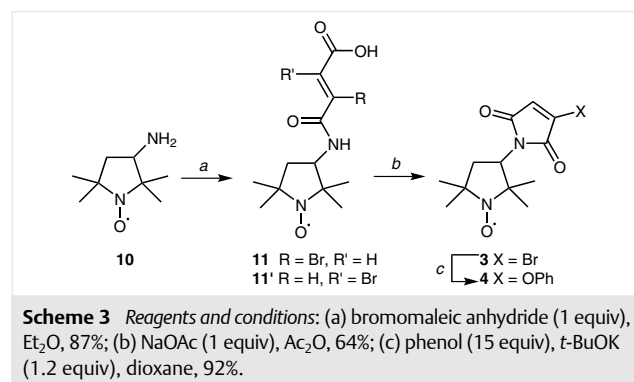
Next-generation maleimides have a leaving group and this can be altered to tune for desired properties. It has previously been shown that bromide may leave so rapidly that adjacent pairs of cysteines will each bind a maleimide marker whereas an *O*-phenyl leaves less efficiently, thus favoring the intramolecular reaction and formation of a succinimide-bridged product.¹⁰ We have therefore made each label type with Br (**1** and **3**) and OPh (**2** and **4**) leaving groups on the maleimide. The two families consist of either a pyrrolinoxyl (**1** and **2**) or pyrrolidinoxyl (**3** and **4**) carbon framework stabilizing the radical. The pyrrolidinoxyl has an extra stereocenter but one less carbon in the linker between the nitroxide ring and the nitrogen of the maleimide. An advantage of the pyrrolinoxyl ring is that it is expected to have greater stabilizing properties for the nitroxyl radical.

Carboxylic acid¹¹ **5** (Scheme 2) was converted into alcohol **7** via the intermediate mixed anhydride **6**. This route was favored over lithium aluminium hydride reduction which can lead to low yield and a lack of reproducibility.¹² The two-step procedure has previously been used by Kirilyuk et al. for the synthesis of 2,5-bis(spirocyclohexane)-substituted nitroxides and in our hands gave **7**, reproducibly, at 83% yield.¹³ Alcohol **7** was then converted into tosylate **8**, and the tosyl group was substituted with ammonia, yielding amine **9**.¹⁴ Efficient amination (80%) was achieved using a 7 N solution of ammonia in methanol. Amine **9** was transformed using a one-step procedure into bromomaleimide **1**

in the presence of bromomaleimic anhydride using acetic acid as the solvent.¹⁵ This reaction gave **1** with 47% yield, and the nitroxyl group was shown to be unaffected by the reaction conditions, that is, no reduction was observed by continuous wave (CW) EPR spectroscopy.



When this one-step bromomaleimide formation reaction was attempted on the commercial aminopyrrolidinoxyl radical **10** the formation of a complex mixture was observed. This mixture did not give any signal in CW EPR spectroscopy, suggesting that reduction of the nitroxyl moiety had occurred. Thus, **3** was synthesized from amine **10**, using a two-step reaction pathway (Scheme 3).¹⁶ This procedure consists of the synthesis of a mixture of compounds **11** and **11'**, using bromomaleic anhydride, followed by an intramolecular cyclization giving compound **3** with an overall yield of 56%.



In both families of NGM labels, the OPh derivatives (**2** and **4**) were synthesized, with an excellent yield, from the bromo precursor (**1** and **3**, respectively) using phenol in the presence of *t*-BuOK.¹⁰

The CW EPR spectra of the compounds did not demonstrate any decrease in radical content over the course of the reactions. The structure of the compounds was confirmed

through NMR spectroscopy, mass spectrometry, and X-ray crystallography. The crystallographic structures are presented in Figure 1. The crystals were racemic and diffracted well to give single independent results for **1**, **2**, and **4**. Crystals of **3** gave less clear results. However, HPLC, TLC, and NMR results for **3** did not indicate that it was less pure than the other end products.

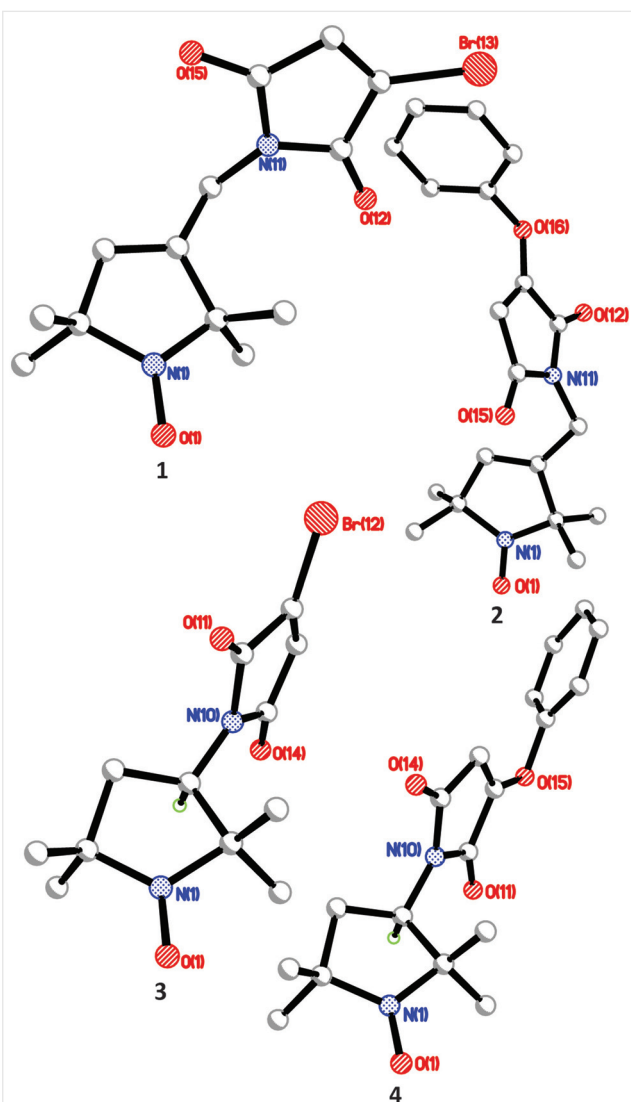


Figure 1 Ball and stick representation of compounds **1–4** from X-ray crystallography. The hydrogen atoms are omitted for clarity except for the stereocenter at C3' for **3** and **4**. For **3** this representation is from one of three crystallographically independent molecules.

To demonstrate that the labels bind to proteins through loss of either their Br or OPh groups, and to see if there were any clear differences in mobility between the families, **1–4** were added to sperm whale myoglobin. This was an available protein containing one cysteine residue for label-

ing at position 3, which is a serine in the wild type.¹⁷ A tenfold excess of spin label was incubated with the protein for one hour at room temperature or at 4 °C overnight. Size-exclusion chromatography was then employed to remove unreacted label. LC-MS confirmed that all four labels bound to myoglobin and gave the expected mass increase, with no free protein observed (Figure 2).¹⁸ The CW EPR spectra for these samples show that the labels have a reduced mobility compared to the free labels which gave a characteristic sharp three-line spectrum (Figure 3). The broader spectra for the shorter linked spin labels **3** and **4** demonstrate that they have a further reduced conformational freedom compared to the longer linkers of **1** and **2**, see Figure 3.^{9,19}

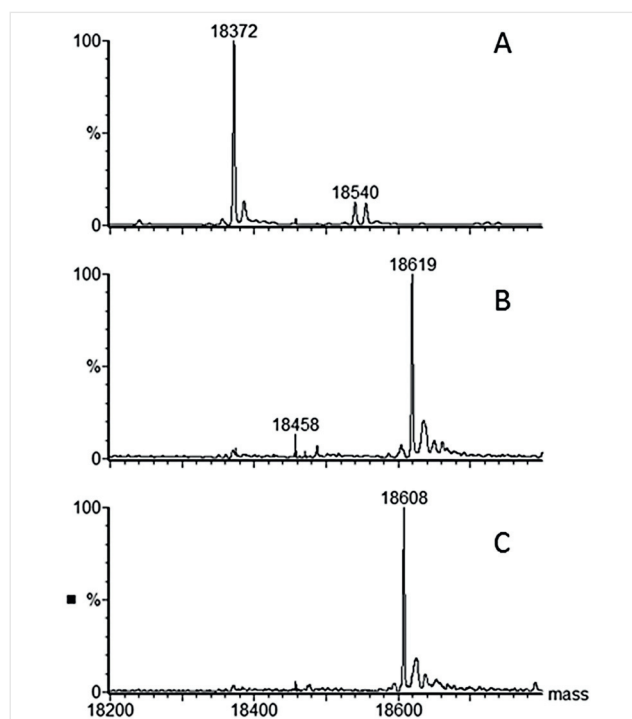


Figure 2 LC-MS spectra showing intact mass of the sperm whale myoglobin S3C (initiator Met present and counted as residue 0). Protein samples were analyzed on a Waters 2795 HPLC and LCT, which had been calibrated with horse heart myoglobin, desalting with a Waters Mass-Prep on-line desalting cartridge, eluting with an increasing gradient of acetonitrile. The envelope of multiply charged signals was deconvoluted using MaxEnt1 software to give the molecular mass of the protein: A) Unlabeled, expected mass 18371.2, found 18372; B) with **2**, following loss of Br[−] expected mass 18620.5, found 18619; C) with **3**, following loss of [−]OPh expected mass 18606.5, found 18608.

In conclusion, four new nitroxide-containing spin labels have been synthesized with next-generation maleimide functionalities. This may lead to the ability to selectively bind the spin label to vicinal cysteines and trials are ongoing.²⁰

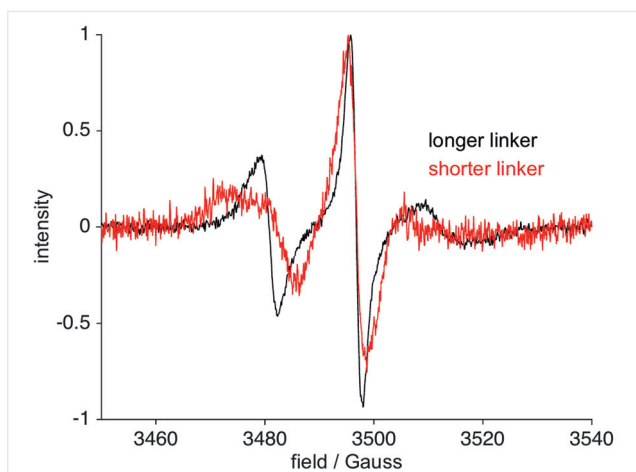


Figure 3 X-band CW EPR spectra at room temperature (Bruker EMX with SHQ resonator) for the labels **2** (longer linker, black) and **4** (shorter linker, red) after addition to sperm whale myoglobin S3C and measured in PBS buffer with 50% glycerol. Data have been overlaid at the central $y = 0$ point to correct for small differences in microwave frequency and have been normalized by their maximum intensity. Spectra were taken under nonsaturating conditions at room temperature with a modulation amplitude of 0.8 G per point.

Acknowledgment

This work was funded by EPSRC, Grant Number EP/LO22044/1, a Royal Society University Research Fellowship to JEL, Royal Society Research Grant RG120645 and support from Active Spectrum Inc and the Wellcome Trust-JIF funding scheme [grant number 063597] which funded the purchase of the Waters 2795 HPLC and LCT mass spectrometer at the BSRC Mass Spectrometry and Proteomics Facility, University of St Andrews. We acknowledge the use of the EPSRC UK National Mass Spectrometry Facility at Swansea University and the MS and NMR services in the EaStCHEM School of Chemistry, St Andrews. We thank Prof. Jason Chin (LMB, Cambridge, UK) for myoglobin plasmid and Prof David O'Hagan for laboratory space in the EaStCHEM School of Chemistry, St Andrews.

Supporting Information

Supporting information for this article is available online at <http://dx.doi.org/10.1055/s-0035-1562451>.

References and Notes

- (1) New address: PharmaCytics Operations BV, Novio Tech Campus, Office 3.12, Transistorweg 5v, 6534 AT Nijmegen, NL.
- (2) New address: Almac Group, 5 Fleming Building, Edinburgh Technopole, Milton Bridge, EH26 0BE, UK.
- (3) Hubbell, W. L.; López, C. J.; Altenbach, C.; Yang, Z. *Curr. Opin. Struct. Biol.* **2013**, *23*, 725.
- (4) Clore, G. M.; Iwahara, J. *Chem. Rev.* **2009**, *109*, 4108.

- (5) (a) Esmann, M.; Sar, P. C.; Hideg, K.; Marsh, D. *Anal. Biochem.* **1993**, *213*, 336. (b) Fawzi, N. L.; Fleissner, M. R.; Anthis, N. J.; Kálai, T.; Hideg, K.; Hubbell, W. L.; Clore, G. M. *J. Biomol. NMR* **2011**, *51*, 105.
- (6) Fiser, A.; Simon, I. *Methods Enzymol.* **2002**, *325*, 10.
- (7) (a) Liu, C. C.; Schultz, P. G. *Annu. Rev. Biochem.* **2010**, *79*, 413. (b) Schmidt, M. J.; Borbas, J.; Drescher, M.; Summerer, D. *J. Am. Chem. Soc.* **2014**, *136*, 1238.
- (8) Smith, M. E. B.; Schumacher, F. F.; Ryan, C. P.; Tedaldi, L. M.; Papaioannou, D.; Waksman, G.; Caddick, S.; Baker, J. R. *J. Am. Chem. Soc.* **2010**, *132*, 1260.
- (9) Schumacher, F. F.; Sanchania, V. A.; Tolder, B.; Wright, Z. V. F.; Ryan, C. P.; Smith, M. E. B.; Ward, J. M.; Caddick, S.; Kay, C. W. M.; Aepli, G.; Chester, K. A.; Baker, J. R. *Sci. Rep.* **2013**, *3*, 1525.
- (10) Marculescu, C.; Kossen, H.; Morgan, R. E.; Mayer, P.; Fletcher, S. A.; Tolner, B.; Chester, K. A.; Jones, L. H.; Baker, J. R. *Chem. Commun.* **2014**, *50*, 7139.
- (11) Rozantsev, E. G. *Free Nitroxyl Radicals*; Plenum Press: New York, **1970**, 203.
- (12) Grimaldi, M.; Scrima, M.; Esposito, C.; Vitiello, G.; Ramunno, A.; Limongelli, V.; D'Errico, G.; Novellino, E.; D'Ursi, A. M. *Biochim. Biophys. Acta* **2010**, *1798*, 660.
- (13) Kirilyuk, I. A.; Polienko, Y. F.; Krumkacheva, O. A.; Strizhakov, R. K.; Gatilov, Y. V.; Grigor'ev, I. A.; Bagryanskaya, E. G. *J. Org. Chem.* **2012**, *77*, 8016.
- (14) Drăguțan, I.; Drăguțan, V.; Melhorn, R. J. *Bull. Soc. Chim. Belg.* **1988**, *97*, 457.
- (15) Martinez-Ariza, G.; Dietrich, J.; De Moliner, F.; Hulme, C. *Synlett* **2013**, *24*, 1801.
- (16) Dreyer, J. L.; Beinert, H.; Keana, J. F.; Hankovszky, O. H.; Hideg, K.; Eaton, S. S.; Eaton, G. R. *Biochim. Biophys. Acta* **1983**, *745*, 229.
- (17) Motion, C. L.; Lovett, J. E.; Bell, S.; Cassidy, S. L.; Cruikshank, P. A. S.; Bolton, D. R.; Hunter, R. I.; El Mkami, H.; Van Doorslaer, S.; Smith, G. M. *J. Phys. Chem. Lett.* **2016**, *7*, 1411.
- (18) The magnetic resonance data supporting this publication can be accessed at <http://dx.doi.org/10.17630/19ac1929-f829-449b-adee-323baeedf0e5>. CCDC 1468783-1468786 contains the supplementary crystallographic data for this paper. The data can be obtained free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/getstructures. The primary data for the LC-MS analysis have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD003824.²¹
- (19) Stone, T. J.; Buckman, T.; Nordio, P. L.; McConnell, H. M. *Proc. Natl. Acad. Sci. U.S.A.* **1965**, *54*, 1010.
- (20) **3-[(Ethoxycarbonyl)oxy]carbonyl]-2,5-dihydro-2,2,5,5-tetramethyl-1H-pyrrol-1-yloxy (6)**
Ethylchloroformate (2.9 g, 26.7 mmol) was added dropwise to a stirred cold (-10°C) solution of **5** (3.6 g, 19.5 mmol) in a mixture of dry toluene (155.0 mL) and Et_3N (3.5 mL, 35.5 mmol). After stirring for 50 min, the solvent was evaporated, and the residue was triturated with Et_2O . The precipitate was filtered off, washed with Et_2O . The organic layer was concentrated in vacuo, and the residue was recrystallized from hexane yielding **6** (5.0 g, quant.) as a yellow solid. ESI-HRMS: m/z calcd for $\text{C}_{12}\text{H}_{18}\text{NO}_5$: 256.1179; found: 256.1176.
3-Hydroxymethyl-2,2,5,5-tetramethylpyrrolidine-N-oxyl (7)
A solution of NaBH_4 (1.5 g, 39.0 mmol) in EtOH was cooled down with an ice bath and **6** (5.0 g, 19.5 mmol) was added portionwise upon stirring. After stirring for 2 h, the solvent was

evaporated under reduced pressure, and the residue was diluted with water and extracted with Et₂O. The extract was dried and concentrated to give **7** (2.8 g, 83%) as a yellow powder. ¹H NMR (400 MHz, acetone *d*₆/D₂O + 1.5 equiv Na₂S₂O₄): δ = 1.14 (s, 6 H), 1.16 (s, 6 H), 4.02 (d, *J* = 1.7 Hz, 2 H), 5.50 (s, 1 H).^{22,23} ESI-HRMS: *m/z* calcd for C₉H₁₆NO₂Na⁺: 193.1073; found: 193.1070.

2,5-Dihydro-2,2,5,5-tetramethyl-3-(((4-methylphenyl)sulfonyl)oxy)methyl)-1H-pyrrol-1-yloxy (8)

A solution of **7** (2.5 g, 14.5 mmol) and triethylamine (1.9 mL, 19.3 mmol) in dry CH₂Cl₂ (40.0 mL) was cooled at –10 °C, and then TsCl (2.8 g, 14.5 mmol) was added portionwise upon vigorous stirring. The solution was stirred for 3 h at r.t., washed with water and with a sat. solution of NaHCO₃, dried, and concentrated under vacuo. The crude was purified by column chromatography (cyclohexane–EtOAc) yielding **8** (3.5 g, 67%) as a yellow solid. ESI-HRMS: *m/z* calcd for C₁₆H₂₂O₄NNa⁺: 347.1162; found: 347.1154.

3-(Aminomethyl)-2,5-dihydro-2,2,5,5-tetramethyl-1H-pyrrol-1-yloxy (9)

A solution of **8** (1.5 g, 4.6 mmol) in anhydrous MeOH was added dropwise into NH₃ solution (75.0 mL, 7 N in MeOH). The mixture was stirred for 2 h at r.t., then left to stand overnight. The solvent was evaporated under reduced pressure. The residue was treated with a buffer solution (60.0 mL; mixture of citric acid and Na₂HPO₄) at pH 5 and extracted with Et₂O. The aqueous layer was saturated with NaOH and extracted with Et₂O. The extract was dried and concentrated yielding **9** (706.0 mg, 80%) as an orange oil. ¹H NMR (400 MHz, CDCl₃ + phenylhydrazine) to reduce the nitroxide radical to a diamagnetic *N*-hydroxylamine,²⁴ since we found that our product was degraded by Na₂S₂O₄: δ = 1.24 (s, 6 H), 1.25 (s, 6 H), 3.31 (d, *J* = 1.8 Hz, 2 H), 5.42 (s, 1 H). ESI-HRMS: *m/z* calcd for C₉H₁₈ON₂⁺: 170.1414; found: 170.1408.

3-[(3-Bromo-2,5-dihydro-1H-pyrrole-2,5-dione-1-yl)-methyl]-2,5-dihydro-2,2,5,5-tetramethyl-1H-pyrrol-1-yloxy (1)

Bromomaleic anhydride (388.2 μL, 4.2 mmol) was dissolved in AcOH (7.0 mL). Nitroxyl **9** (707.9 mg, 4.2 mmol) in AcOH (7.0 mL) was added, and the reaction was heated at 80 °C for 3 h. The solvent was removed under vacuo, and the mixture was purified by column chromatography (cyclohexane–EtOAc) to give the bromomaleimides **1** (644.7 mg, 47%) as an orange powder; mp 146–147 °C. ¹H NMR (400 MHz, CDCl₃ + phenylhydrazine): δ = 1.38 (s, 6 H), 1.45 (s, 6 H), 4.17 (d, *J* = 1.4 Hz, 2 H), 5.40 (s, 1 H), 6.94 (s, 1 H). NSI-HRMS: *m/z* calcd for C₁₃H₁₇O₃N₂Br⁺: 328.0417; found: 328.0417.

3-(3-Bromo-2,5-dihydro-1H-pyrrole-2,5-dione-1-yl)-2,5-dihydro-2,2,5,5-tetramethyl-1H-pyrrol-1-yloxy (3)

To a stirred solution of amine **10** (489.0 mg, 3.1 mmol) in dry

Et₂O (24.7 mL) bromomaleic anhydride (286.7 μL, 3.1 mmol) was added. The reaction was left stirring at r.t. for 3 h. The precipitate was filtered and washed with Et₂O yielding a mixture of **11** and **11'** (910.0 mg, 87%) as a yellow solid.

The mixture of **11** and **11'** (910 mg, 2.7 mmol) and NaOAc (222.2 mg, 2.7 mmol) was dissolved in Ac₂O (13.5 mL) and heated at 60–70 °C for 3 h. The reaction mixture was then concentrated, dissolved in CH₂Cl₂ and filtered. The filtrate was concentrated and purified by column chromatography (cyclohexane–EtOAc) to give **3** (853 mg, 64%) as an orange solid; mp 101–102 °C. ¹H NMR (400 MHz, CDCl₃ + phenylhydrazine): δ = 1.08 (s, 3 H), 1.25 (s, 3 H), 1.26 (s, 3 H), 1.36 (s, 3 H), 1.82 (dd, *J* = 12.5, 8.8 Hz, 1 H), 2.93 (dd, *J* = 12.5, 11.0 Hz, 1 H), 4.47 (dd, *J* = 11.0, 8.8 Hz, 1 H), 6.89 (s, 1 H). ESI-HRMS: *m/z* calcd for C₁₂H₁₆O₃N₂BrNa⁺: 338.0237; found: 338.0223.

Phenoxymaleimide 2 and 4 – General Procedure

To molten phenol (13.2 mmol), *t*-BuOK (1.1 mmol) in dry dioxane (0.8 mL) was added dropwise, and the solution was left stirring for 10 min at 40 °C. Then a solution of bromomaleimide **1** or **3** (0.8 mmol) in dry dioxane (0.8 mL) was added dropwise, and the resulting mixture was stirred at 40 °C for 30 min. After this time, the solvent was evaporated under reduced pressure. The mixture was purified by column chromatography (cyclohexane–EtOAc) to give the corresponding phenoxymaleimide.

3-[(3-Phenoxy-2,5-dihydro-1H-pyrrole-2,5-dione-1-yl)-methyl]-2,5-dihydro-2,2,5,5-tetramethyl-1H-pyrrol-1-yloxy (2)

Yield 279 mg (92%); light yellow powder; mp 92–94 °C. ¹H NMR (400 MHz, CDCl₃ + phenylhydrazine): δ = 1.32 (s, 6 H), 1.40 (s, 6 H), 4.14 (d, *J* = 1.4 Hz, 2 H), 5.34 (s, 1 H), 5.48 (s, 1 H), 7.36–7.29 (m, 2 H), 7.51–7.43 (m, 3 H). NSI-HRMS: *m/z* calcd for C₁₉H₂₂O₄N₂⁺: 342.1574; found: 342.1570.

3-[(3-Phenoxy-2,5-dihydro-1H-pyrrole-2,5-dione-1-yl)-2,5-dihydro-2,2,5,5-tetramethyl-1H-pyrrol-1-yloxy (4)

Yield 345 mg (92%); yellow solid; mp 120–122 °C. ¹H NMR (400 MHz, CDCl₃ + phenylhydrazine): δ = 1.13 (s, 3 H), 1.24 (s, 3 H), 1.27 (s, 3 H), 1.36 (s, 3 H), 1.80 (dd, *J* = 12.4, 8.7 Hz, 1 H), 2.98 (dd, *J* = 11.2, 12.4 Hz, 1 H), 4.46 (dd, *J* = 11.2, 8.7 Hz, 1 H), 5.29 (s, 1 H), 7.37–7.30 (m, 2 H), 7.54–7.43 (m, 3 H). ESI-HRMS: *m/z* calcd for C₁₈H₂₁N₂O₄Na⁺: 352.1394; found: 352.1371.

- (21) Vizcaíno, J. A.; Csordas, A.; del-Toro, N.; Dienes, J. A.; Griss, J.; Lavidas, I.; Mayer, G.; Perez-Riverol, Y.; Reisinger, F.; Ternent, T.; Xu, Q. W.; Wang, R.; Hermjakob, H. *Nucleic Acids Res.* **2016**, 44, 01 D447.
- (22) Hideg, K.; Hankovszky, H. O.; Lex, L.; Kulcsár, G. *Synthesis* **1980**, 911.
- (23) Powell, J. H.; Johnson, E. M. II.; Gannett, P. M. *Molecules* **2000**, 5, 1244.
- (24) Lee, T. D.; Keana, J. F. W. *J. Org. Chem.* **1975**, 40, 3145.