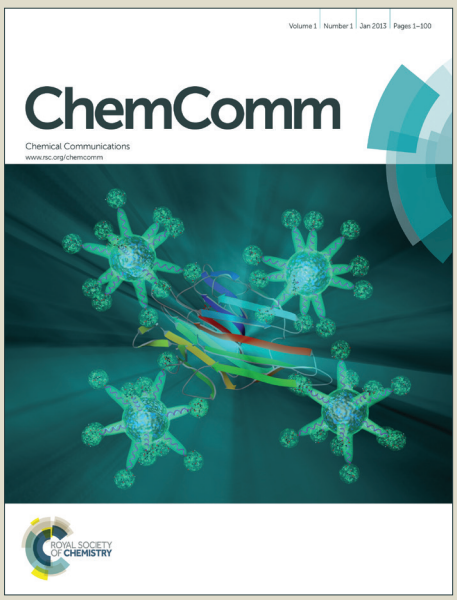
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

The nitrilimine-alkene cycloaddition is an ultra rapid click reaction

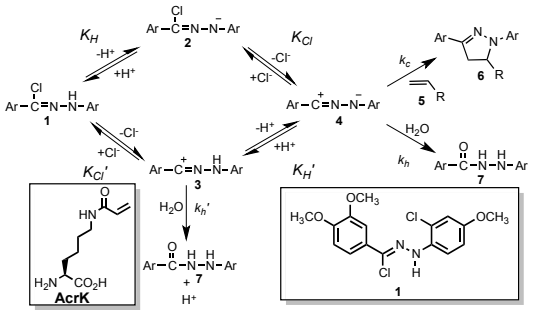
Xiaoshan Shayna Wang,[a] Yan-Jiun Lee[a] and Wenshe R. Liu[a]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
DOI: 10.1039/b000000x

The transient formation of nitrilimine in aqueous conditions is greatly influenced by pH and chloride. At a basic condition (pH 10) with no chloride, a diarylnitrilimine precursor readily ionizes to form diarylnitrilimine that reacts almost instantaneously with an acrylamide-containing protein and fluorescently labels it.

A recent mining of organic reactions for click labeling of proteins has revamped tetrazine-based Diels-Alder[1-5] and cyclooctyne-based 1,3-dipolar cycloadditions.[6-9] Unlike the Cu^{+}-catalysed click reaction,[10-12] both tetrazine and cyclooctyne cycloadditions undergo spontaneously in aqueous conditions, avoiding side reactions potentially induced by a transition metal catalyst.[13] A noteworthy advantage of tetrazine-based click reaction is its fast reaction kinetics. A hitherto fastest reported tetrazine-transcyclooctene reaction has a second-order k as 2.8×10^{6} $M^{-1}s^{-1}$.[14] Cyclooctyne was originally explored for its spontaneous reaction with azide and recently extended to react with nitrone and tetrazine.[6, 15, 16] Cyclooctyne reacts rapidly with tetrazine.[17] Derivatives of cyclooctyne that react with azide rapidly and nitrone with a second-order k close to 50 $M^{-1}s^{-1}$ have also been developed.[18-20] Another copper-free click reaction that has been recently explored but not yet highly appreciated is the nitrilimine-alkene cycloaddition.[21] On contrary to tetrazine and cyclooctyne that stably exist in aqueous conditions, nitrilimine reacts with water, therefore needs to be formed transiently.[22] Two methods are generally used to transiently form nitrilimine. One is the photolysis of tetrazole and the other is the ionization of hydrazonyl halide.[21, 23] Lin *et al.* have extended the first approach for photoclick protein labeling in living cells.[24-26] The second approach was recently explored to undergo fluorescent turn-on labeling of proteins incorporated with norbornene, cyclopropene, and acrylamide moieties.[22, 27, 28] Reaction kinetics of the nitrilimine-alkene cycloaddition that involved tetrazole and hydrazonyl chloride as nitrilimine precursors were previously characterized.[22, 24] All these characterizations were performed in a PBS-acetonitrile (1:1) buffer. The high concentration of chloride (140 mM) in PBS potentially offsets the transient formation of nitrilimine and consequently curbs its reaction with alkene. Here we report a comprehensive study of pH and chloride dependences of the nitrilimine-alkene cycloaddition reaction and demonstrate that it is an ultra rapid click reaction for protein labeling at a basic condition (pH 10) with no chloride.

We chose a hydrazonyl chloride (**1** in **Scheme 1**) as a nitrilimine precursor for our kinetic analysis due to the difficulty of quantitative photolysis of a tetrazole to form a nitrilimine that nonetheless reacts with water and chloride in aqueous conditions to form a hydrazonyl chloride. In an aqueous buffer with a high chloride concentration, **1** presumably undergoes two parallel ionization processes to lose a proton and a chloride to generate a nitrilimine **4**. **4** then reacts either with **5** to form a fluorescent cycloaddition product **6** or with water to produce **7**. **7** can also be formed from the hydrolysis of **3**. These two parallel ionization processes of hydrazonyl halide in aqueous conditions were studied and demonstrated previously.[29] Assuming the ionization of **1** to **4** is a fast equilibrium, the **6** formation will follow eq. 1 ($[\mathbf{6}] = \frac{k_c\cdot[\mathbf{5}]}{k_c\cdot[\mathbf{5}]+k_h+k_h'\cdot\frac{[H^+]}{K_{H'}}}\cdot[\mathbf{1}]_0\cdot\left(1-e^{-\left(k_c\cdot\frac{K_H\cdot K_{Cl}}{[H^+]\cdot[Cl^-]}[\mathbf{5}]+k_h\cdot\frac{K_H\cdot K_{Cl}}{[H^+]\cdot[Cl^-]}+k_h'\cdot\frac{K_{Cl}'}{[Cl^-]}\right)t}\right)$) when **5** is excessive, $[H^+]>>K_H$, and $[Cl^-]>>K_{Cl}$. In eq. 1, $[\mathbf{1}]_0$ represents the initial concentration of **1**. Based on eq. 1, the pseudo first-order k of the **6** formation can be described as eq. 2 ($k_{app} = k_c\cdot\frac{K_H\cdot K_{Cl}}{[H^+]\cdot[Cl^-]}\cdot[\mathbf{5}] + k_h\cdot\frac{K_H\cdot K_{Cl}}{[H^+]\cdot[Cl^-]} + k_h'\cdot\frac{K_{Cl}'}{[Cl^-]}$). This equation can be further simplified as eq. 3 ($k_{app} = k_{c(obs)}\cdot[\mathbf{5}] + k_{h(obs)}$) where $k_{c(obs)} = k_c\cdot\frac{K_H\cdot K_{Cl}}{[H^+]\cdot[Cl^-]}$ (eq. 4). Since **6** is highly fluorescent, its formation can be facilely detected using a fluorescent spectrophotometer and analysed to obtain k_{app}. The determined k_{app} values at varied concentrations of **5** at a given pH and a chloride concentration can be applied to obtain $k_{c(obs)}$. In principle, the determined $k_{c(obs)}$ values at varying pH and chloride concentrations will allow to assess k_c, the second-order k of the nitrilimine-alkene cycloaddition in aqueous conditions.



Scheme 1: The nitrilimine-alkene reaction in the presence of chloride.

We first studied the pH dependence of $k_{c(obs)}$ at 50 mM chloride. Reactions between 5 μM **1** and varying concentrations of acrylamide at five given pH values (6-10) were monitored using a PTI QM-40 fluorescent spectrophotometer with an excitation light at 320 nm and an emission detection at 480 nm. The fluorescent increment data were fitted to a single exponential increase equation to obtain k_{app}. The resolved k_{app} values were then plotted against the acrylamide concentrations. As shown in the inset of **Figure 1A,** k_{app} is linearly dependent on the acrylamide concentration at a given pH and the data were readily used to determine $k_{c(obs)}$, validating the mechanism proposed in **Scheme 1**. Although $\log(k_{c(obs)})$ shows a linear dependence on pH (**Figure 1A**) as eq. 4 predicts (eq. 4 can be transformed as $log(k_{c(obs)}) = log\left(\frac{k_c\cdot K_H\cdot K_{Cl}}{[Cl^-]}\right) + pH$), the data can not be simply

fitted to eq. 4. They are best fitted to eq. 5 ($k_{c(obs)} = k_c \cdot \frac{(K_H)^x \cdot K_{Cl}}{[H^+]^x \cdot [Cl^-]}$) with an x value as 0.64±0.01 and $k_c \cdot \frac{(K_H)^{0.64} \cdot K_{Cl}}{[Cl^-]}$ as (1.42±0.01)×10^{-5}. A deviation from eq. 4 may be due to the presence of chloride that changes the proton activity during the ionization process. This is highly possible since a similar deviation was not observed for reactions in conditions without chloride, which will be presented later. **Figure 1A** clearly shows that the observed cycloaddition rate constant increased about 200 fold when pH was changed from 6 to 10. Therefore, when an acrylamide-containing protein is labeled with **1** at different pH, faster labeling rates are expected at higher pH values. To approve this, we performed the labeling of sfGFP2AcrK (a superfolder green fluorescent protein with N^{ε}-acryloyl-lysine (AcrK in **Scheme 1**) incorporated at its S2 position) by 150 μM **1** for 15 min at 50 mM chloride and pH from 6 to 10. The expression of sfGFP2AcrK followed a method described previously.[22] Presented in **Figure 1B**, the labeling efficiency is clearly pH dependent, with higher pH leading to more efficient labeling.

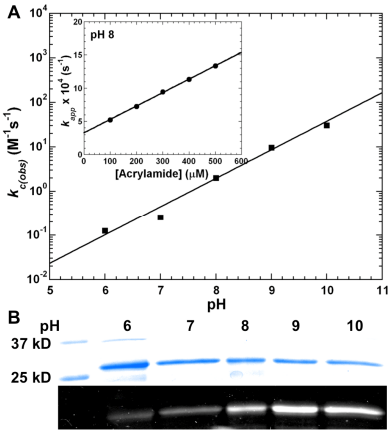


Figure 1: (**A**) the pH dependence of $k_{c(obs)}$. The inset shows the acrylamide concentration dependence of k_{app} at pH 8 and 50 mM chloride in acetonitrile-50 mM phosphate buffer (1:1). (**B**) The labeling efficiency of sfGFP2AcrK by **1** at different pH. The labeling reactions between 5 μM sfGFP2AcrK and 150 μM **1** were carried out in acetonitrile-50 mM phosphate buffer (1:1) for 15 min before 500 mM acrylamide was added to sequestrate **1** from reacting with sfGFP2AcrK and then the labeling solutions were analyzed by SDS-PAGE. The top panel shows the Coomassie blue stained gel and the bottom panel presents a fluorescent image of the same gel before it was stained by Coomassie blue.

Eq. 5 also indicates an inverse linear dependence of k_{app} on the chloride concentration, which has been approved by our kinetic analyses performed in varying chloride concentrations and pH 9. At a particular chloride concentration (10~100 mM), the determined k_{app} values are linearly dependent on the acrylamide concentrations, which were used to obtain $k_{c(obs)}$. Plotting $k_{c(obs)}$ against 1/[Cl$^-$] indeed shows a linear dependence (**Figure 2A**). We also did similar kinetic analyses at 1 mM chloride. Although the determined k_{app} values are much higher than those determined at higher chloride concentrations, k_{app} values at different acrylamide concentrations are almost constant, and therefore not valid to calculate $k_{c(obs)}$. It is possible that at a low chloride concentration the two dechlorination processes (**2** to **4** and **1** to **3**) do not reach fast equilibria, invalidating **Scheme 1** and eq. 1 in data analysis. This study clearly shows a strong inhibitory effect of chloride on the nitrilimine-alkene cycloaddition, indicating that applying the nitrilimine-alkene cycloaddition for protein labeling needs to avoid a high chloride concentration. This is exactly what we observed in labeling sfGFP2AcrK with **1** at pH 7 and different chloride concentrations (**Figure 2B**). A 30 min labeling reaction in the absence of chloride led to an intensely fluorescently labeled protein. The labeling efficiency gradually diminished to barely detectable when chloride was increased from 0 to 200 mM.

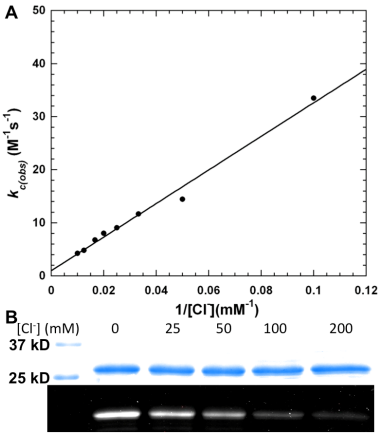
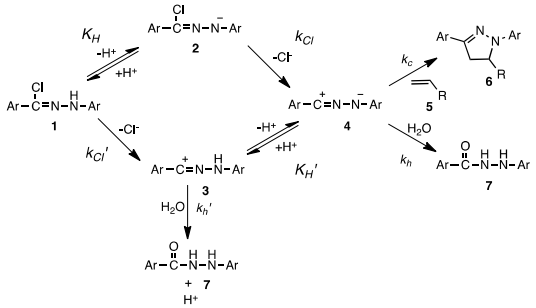


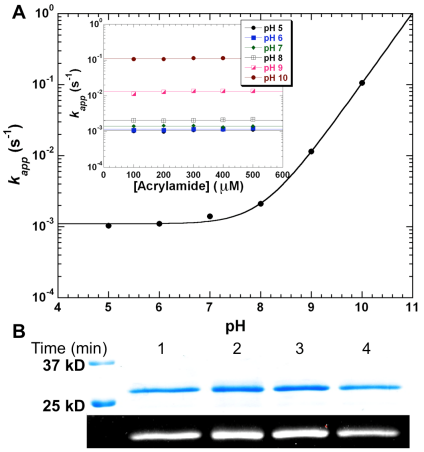
Figure 2: (**A**) the chloride dependence of $k_{c(obs)}$. (**B**) The labeling efficiency of sfGFP2AcrK by **1** at pH 7 and different chloride concentrations. The labeling reactions between 5 μM sfGFP2AcrK and 150 μM **1** were carried out in acetonitrile-50 mM phosphate buffer(1:1), pH 7, and varying chloride concentrations for 30 min before adding 500 mM acrylamide and then SDS-PAGE analysis (top: Coomassie blue stained; bottom: fluorescent).



Scheme 2: The nitrilimine-alkene reaction in the absence of chloride

The aforementioned kinetic analyses at 1 mM chloride prompted us to look into the reaction kinetics of the nitrilimine-alkene cycloaddition in the absence of chloride. Without chloride, the ionization of **1** and subsequent reactions with water and alkene will presumably follow a mechanism presented in **Scheme 2**. The two dechlorination steps become rate limiting. Since chloride released from **1** at a concentration much lower than 1 mM will not eviscerate the mechanism shown in **Scheme 2**. **Scheme 2** ensues a formation of **6** following eq. 6 ($[\mathbf{6}] = \frac{k_c \cdot [5]}{k_c \cdot [5] + k_h + k_{h'} \cdot \frac{[H^+]}{K_{H'}}} \cdot [\mathbf{1}]_0 \cdot (1 - e^{-\left(k_{Cl} \cdot \frac{K_H}{[H^+]} + k_{Cl'}\right)t})$) that results in an apparent rate constant defined as eq. 7 ($k_{app} = k_{Cl} \cdot \frac{K_H}{[H^+]} + k_{Cl'}$). Eq. 7 shows that k_{app} is inversely proportional to the proton concentration but not related to the provided acrylamide, which was observed during our kinetic analyses of the nitrilimine-alkene cycloaddition in the absence of chloride. When reactions between 5 μM **1** and different acrylamide concentrations were performed at a specific pH without chloride, all resulted in a same reaction rate constant (inset of **Figure 3A**). Raising pH led to higher reaction rate constants. The logarithms of determined k_{app} values as a function of pH are presented in **Figure 3A**, which can be well fitted to eq. 7. At pH 10 with no chloride, the determined k_{app} is 0.111±0.002 s^{-1}. Since this rate constant is not related to the concentrations of both **1** and acrylamide, using **1** to label a protein with an acrylamide moiety at any concentrations of **1** and the protein will have a labeling half life close to 6 s when the chloride anion is absent in labeling conditions, achieving almost

instantaneous protein labeling. To demonstrate this rapid labeling process, we tested the labeling of sfGFP2AcrK by **1** at pH 10 for different lapses of time. As shown in **Figure 3B**, labeling



sfGFP2AcrK with **1** for 1 to 4 min all led to an intensely fluorescently labeled protein with equally fluorescent intensities, implying labeling was mostly completed within 1 min.

Figure 3: (**A**) the pH dependence of k_{app} in the absence of chloride. The data were fitted to eq. 7. The inset shows the acrylamide concentration dependence of k_{app} at pH 5-10 in acetonitrile-50 mM phosphate buffer(1:1) without chloride. (**B**) The labeling efficiency of sfGFP2AcrK by 1 at pH 10 without chloride. The labeling reactions between 5 μM sfGFP2AcrK and 150 μM **1** were carried out in acetonitrile-50mM phosphate buffer(1:1) without chloride for different lapses of time (1-4 min) before adding 500 mM acrylamide and then SDS-PAGE analysis (top: Coomassie blue stained; bottom: fluorescent).

Being a catalyst-free click reaction type, the nitrilimine-alkene cycloaddition has been explored for click and photo-click labeling of proteins. All previously kinetic characterizations of the nitrilimine-alkene cycloaddition were completed in PBS buffers. The current study clearly shows that all previously measured second-order *k*'s of the nitrilimine-alkene cycloaddition are apparent second-order *k*'s that are significantly influenced by pH and chloride. Based on eq. 5, one would need to determine K_H and K_{Cl} to calculate k_c. When we derived eq. 1, we put preconditions that are $[H^+]>>K_H$ and $[Cl^-]>>K_{Cl}$. When the conditions $[H^+]>>K_H$ and $[Cl^-]>>K_{Cl}$ do not hold, the determined apparent k will in theory follow eq. 8 ($k_{c(obs)} = k_c \cdot \frac{K_H \cdot K_{Cl}}{([H^+]+K_H)\cdot([Cl^-]+K_{Cl})}$) but should be best described as eq. 9 ($k_{c(obs)} = k_c \cdot \frac{(K_H)^{0.64}\cdot K_{Cl}}{([H^+]+K_H)^{0.64}\cdot([Cl^-]+K_{Cl})}$) due to the proton activity deviation from what is indicated by pH. As indicated by eq. 9, in a specific chloride concentration, $k_{c(obs)}$ will reach to a plateau when $[H^+]<<K_H$. Since we did not observe the trend of $k_{c(obs)}$ to become saturated to pH 10, a safe guess of a K_H value is small than 10^{-12}. Similarly, $k_{c(obs)}$ showed an inverse proportional dependence of the chloride concentration to lower than 10 mM. A safe estimate of a K_{Cl} value is small than 10^{-3}. We have determined that at 50 mM chloride $k_c \cdot \frac{(K_H)^{0.64}\cdot K_{Cl}}{[Cl^-]}$ is 1.42×10^{-5}. With two estimated values of K_H and K_{Cl}, we can easily determine a k_c value higher than 3.4×10^4 $M^{-1}s^{-1}$. This rate constant is comparable to that of the rapid transcyclooctene-tetrazine cycloaddition and makes the nitrilimine-alkene cycloaddition as one of the fastest click reactions. Another implication of our study is different labeling kinetics in extracellular and intracellular spaces when the nitrilimine-alkene reaction is applied for *in vivo* labeling. Mammalian cells maintain intracellular chloride concentration much lower than their extracellular environments.[30] This large chloride concentration variation may allow to apply the nitrilimine-alkene reaction to specifically achieve intracellular protein sensitization.

Notes and references

[a]Department of Chemistry, Texas A&M University, College Station, TX 77843; wliu@chem.tamu.edu.

Electronic Supplementary Information (ESI) available: kinetic analysis, protein expression, protein labeling, and equation derivation. See DOI: 10.1039/c000000x/

1. M. L. Blackman, M. Royzen and J. M. Fox, *J Am Chem Soc*, 2008, **130**, 13518-13519.
2. N. K. Devaraj, R. Upadhyay, J. B. Haun, S. A. Hilderbrand and R. Weissleder, *Angew Chem Int Ed Engl*, 2009, **48**, 7013-7016.
3. N. K. Devaraj, S. Hilderbrand, R. Upadhyay, R. Mazitschek and R. Weissleder, *Angew Chem Int Ed Engl*, 2010, **49**, 2869-2872.
4. M. T. Taylor, M. L. Blackman, O. Dmitrenko and J. M. Fox, *J Am Chem Soc*, 2011, **133**, 9646-9649.
5. D. S. Liu, A. Tangpeerachaikul, R. Selvaraj, M. T. Taylor, J. M. Fox and A. Y. Ting, *J Am Chem Soc*, 2012, **134**, 792-795.
6. J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli and C. R. Bertozzi, *Proc Natl Acad Sci U S A*, 2007, **104**, 16793-16797.
7. J. A. Codelli, J. M. Baskin, N. J. Agard and C. R. Bertozzi, *J Am Chem Soc*, 2008, **130**, 11486-11493.
8. P. V. Chang, J. A. Prescher, E. M. Sletten, J. M. Baskin, I. A. Miller, N. J. Agard, A. Lo and C. R. Bertozzi, *Proc Natl Acad Sci U S A*, 2010, **107**, 1821-1826.
9. E. M. Sletten, H. Nakamura, J. C. Jewett and C. R. Bertozzi, *J Am Chem Soc*, 2010, **132**, 11799-11805.
10. H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew Chem Int Ed Engl*, 2001, **40**, 2004-2021.
11. V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew Chem Int Ed Engl*, 2002, **41**, 2596-2599.
12. T. R. Chan, R. Hilgraf, K. B. Sharpless and V. V. Fokin, *Org Lett*, 2004, **6**, 2853-2855.
13. B. Wu, Z. Wang, Y. Huang and W. R. Liu, *Chembiochem*, 2012, **13**, 1405-1408.
14. R. Rossin, S. M. van den Bosch, W. Ten Hoeve, M. Carvelli, R. M. Versteegen, J. Lub and M. S. Robillard, *Bioconjug Chem*, 2013, **24**, 1210-1217.
15. X. Ning, R. P. Temming, J. Dommerholt, J. Guo, D. B. Ania, M. F. Debets, M. A. Wolfert, G. J. Boons and F. L. van Delft, *Angew Chem Int Ed Engl*, 2010, **49**, 3065-3068.
16. T. Plass, S. Milles, C. Koehler, J. Szymanski, R. Mueller, M. Wiessler, C. Schultz and E. A. Lemke, *Angew Chem Int Ed Engl*, 2012, **51**, 4166-4170.
17. K. Lang, L. Davis, S. Wallace, M. Mahesh, D. J. Cox, M. L. Blackman, J. M. Fox and J. W. Chin, *J Am Chem Soc*, 2012, **134**, 10317-10320.
18. C. G. Gordon, J. L. Mackey, J. C. Jewett, E. M. Sletten, K. N. Houk and C. R. Bertozzi, *J Am Chem Soc*, 2012, **134**, 9199-9208.
19. C. S. McKay, J. Moran and J. P. Pezacki, *Chem Commun*, 2010, **46**, 931-933.
20. C. S. McKay, M. Chigrinova, J. A. Blake and J. P. Pezacki, *Org Biomol Chem*, 2012, **10**, 3066-3070.
21. R. Huisgen, *Angew Chem Int Ed Engl*, 1963, **2**, 565-598.
22. Y. J. Lee, B. Wu, J. E. Raymond, Y. Zeng, X. Fang, K. L. Wooley and W. R. Liu, *ACS chem biol*, 2013, **8**, 1664-1670.
23. R. Huisgen, R. Grashey, M. Seidel, R. Schmidt, H. Knupfer and G. Wallbillich, *Annalen Der Chemie-Justus Liebig*, 1962, **653**, 105-113.
24. Y. Wang, W. Song, W. J. Hu and Q. Lin, *Angew Chem Int Ed Engl*, 2009, **48**, 5330-5333.
25. W. Song, Y. Wang, J. Qu and Q. Lin, *J Am Chem Soc*, 2008, **130**, 9654-9655.
26. W. Song, Y. Wang, J. Qu, M. M. Madden and Q. Lin, *Angew Chem Int Ed Engl*, 2008, **47**, 2832-2835.
27. E. Kaya, M. Vrabel, C. Deiml, S. Prill, V. S. Fluxa and T. Carell, *Angew Chem Int Ed Engl*, 2012, **51**, 4466-4469.
28. Z. Yu, Y. Pan, Z. Wang, J. Wang and Q. Lin, *Angew Chem Int Ed Engl*, 2012, **51**, 10600-10604.
29. A. F. Hegarty, M. P. Cashman and F. L. Scott, *J. Chem. Soc., Perkin Trans.*, 1972, **2**, 44-52.
30. D. Billups and D. Attwell, *J Physiol*, 2002, **545**, 183-198.