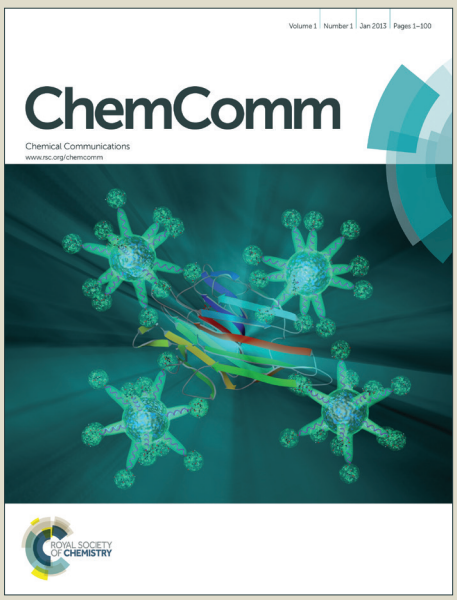
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

RSCPublishing

COMMUNICATION

A Phthalimidation Protocol That Follows Protein Defined Parameters

Cite this: DOI: 10.1039/xoxxoooooox

Rohith Singudas,[a]† Srinivasa Rao Adusumalli,[a]† Pralhad Namdev Joshi,[a] and Vishal Rai[a]*

Received ooth January 2012,
Accepted ooth January 2012

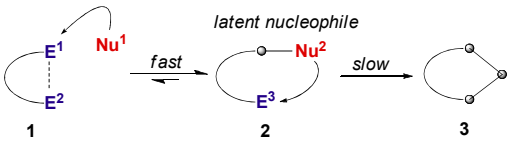
DOI: 10.1039/xoxxoooooox

www.rsc.org/

This work outlines the first phthalimidation protocol suitable for protein labeling and performed in aqueous medium at room temperature and neutral pH with no obligation of catalyst or co-reagent. The methodology is suitable for a range of amines and its efficiency was determined with chemoselective and site-selective labeling of a protein.

Organic reactions with small molecules have evolved over generations and provide access to enormous library of compounds through simple chemical transformations. Yet, chemical modification of protein residues have posed stark challenges owing to unique set of protein defined parameters. Change in the landscape of reaction conditions such as solvent, substrate concentration, temperature and pH has proved non-trivial.[1] Adhering to the redefined conditions and achieving high reaction rate constant has also been arduous. Additional definite challenge has been associated to the low chemoselectivity of reactants or co-reactants. For example, azides can react with cysteine thiols,[2] and phosphines can reduce disulfide bonds and result heterogeneous protein milieu in Staudinger ligation.[3] In a recent report for a protein N-terminal modification under oxidative conditions, cysteine thiols undergo side reactions.[4] Under these circumstances, only a few methodologies qualify and find their application in protein conjugation.[5] Such chemical tools for mild and site-specific derivatization of proteins is desired for their wide range of applications in understanding biological interactions,[6] ligand discovery,[7] disease diagnosis,[8] and high-throughput screening.[9] These transformations can also route access to the knowledge of encoded protein function and enhance genetic information pool.[10] Typically, these indispensable diagnostic tools can be accessed by the derivatization of nucleophilic side chains in case of unmodified proteins.[11] However, the diversity of reactions that can be performed with these functional groups in physiological conditions is very limited.[12] This is largely due to the restrictions over major synthetic chemist's tools such as catalyst, co-catalyst and reaction conditions. The task is particularly challenging because control over reactivity of substrate is the only preferred parameter for modulation. One of the successful example in this perspective is showcased with Cu free click chemistry enabled by strained cyclooctynes.[13]

En route to development of a convenient methodology for protein conjugation, it is imperative to overcome the barrier imposed by low reaction rate of a chemical transformation at working concentration of proteins. This necessitates the use of 10-1000 folds excess reagent to render practically useful conversions. Our primary focus was to enable the transformation in physiological conditions in the absence of a catalyst or an interfering co-reactant. We hypothesized that a two-step one-pot process – initiated by a fast reversible intermolecular step forming an amphoteric intermediate **2** capable of rendering rate determining irreversible intramolecular step (Scheme 1), can prove critical.

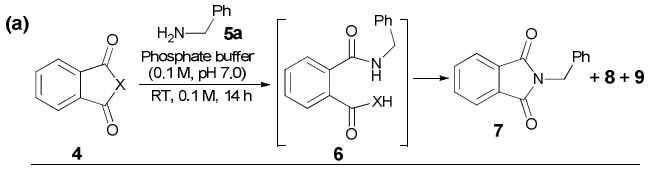


Scheme 1. Generation of an amphoteric intermediate.

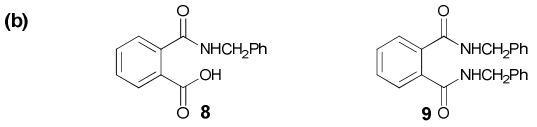
Phthalic anhydride or pthalimides were identified for the generation of amphoteric intermediate. In retrospect, the imide derivatives are important functionalities and present number of applications in chemistry,[14] polymers,[15] and chemical biology.[16] Additionally, phthalimides have served as solvatochromic fluorophore in detection of biomolecular interactions.[17] However, all the reported methods drive on factors such as high temperature, strong bases, and coupling reagents.[14] A biomolecule compatible pthalimidation protocol has remained elusive till day. Here, we report phthalimidation of primary amines in aqueous buffer at ambient temperature, physiological pH in typically 4-48 hours. The success of transformation holds in designing the substrate for appropriate reactivity of the amphoteric intermediate. The methodology proved efficient for labeling of RNase A. The chemoselectivity is inherent to the methodology, whereas the site-selectivity is driven by kinetic preferences of amino acid residue at N-terminus and rendered K1 monolabeled RNase A. The phthalimidoamine product offers an electrophilic center suitable for secondary amine catalyzed late-stage modification. The reported protocol overcomes the drawback of instability of reagent,

intermediate or product towards hydrolysis associated with several other bioconjugation methodologies based on lysine modification.[12]

In a model reaction, benzylamine was stirred with phthalic anhydride **4a** and phthalimides **4b-e** for its efficiency to undergo transimidation and give phthalimidoamines (Scheme 2). The reaction with **4a** results in intermediate **6** (64%) with high threshold for conversion to product **7** (entry 1, Scheme 2). Phthalimide **4b** resulted in 67% conversion and extended reaction time offered no improvement. Diphenylphosphate derivative of phthalimide **4c** was unreactive, whereas N-allyloxyphthalimide **4d** gave a mixture of intermediate **6d**, product **7** and unreacted starting material **4d** after 14 h. N-hydroxyphthalimide **4e** reacted with benzyl amine to give the product **7** in >99% conversions.



Entry	4, X	Product (% Conversion)
1	**4a**, O	**7** (4%), **6a** (64%), **4a** (32%)
2	**4b**, NH	**7** (67%), **4b** (33%)
3	**4c**, N–P(O)(OPh)(OPh)	**7** (<5%), **4c** (>95%)
4	**4d**, N–O–allyl	**7** (67%), **6d** (10%), **4d** (21%)
5	**4e**, NOH	**7** (>99%)



Side products: **8** with **4b** and **4c** in >24 h; **9** with **4e** with >10 equiv. **5a**

Scheme 2. (a) Evaluation of reagents for pthalimidation, (b) side products **8** and **9** in phthalimidation.

The reaction involves a one-pot two-stage transformation of N-hydroxyphthalimide (NHP) **4e** with primary amine **5a** involving a ring opening reaction followed by transimidation to provide phthalimide labeled primary amine **7** (Figure 1). The intramolecular nucleophilic addition of latent nucleophile (see intermediate **6e**) to electrophilic carbonyl is the rate determining step (rds). Design of reagent **4** plays a crucial role in success of this transformation as Z indirectly controls the electrophilicity of carbonyl by regulating n to π* contributions (see second order perturbation theory analysis in ESI for details).

Various aliphatic and aromatic amines with disparate nucleophilicity were tested for their reactivity with NHP **4e** (Scheme 3a). The transformation was efficient with range of aliphatic and aromatic amines and products **10a-10n** were formed in 40-85% yield (also see Table S1, ESI). NHP is partially soluble in phosphate buffer at pH 7.0. A common observation includes disappearance of NHP within 30-120 seconds as it converts to the intermediate **6**. A clear light yellow solution stirred further results in precipitated product that can be filtered and triturated to render the analytically pure product. Filtering the product gradually with time shows no effect on the reaction time. There were few exceptions where amines were insoluble in buffer e.g. tryptamine (product **10d**). For such cases, amines were dissolved in DMSO and dispensed in reaction mixture (10% DMSO in buffer). The products soluble in water, e.g. **10g**, were isolated by liquid-liquid extraction. Reaction with aminoacetaldehyde diethyl acetal showcases the importance of mild reaction conditions where the product **10h** was isolated without tracing any side product.

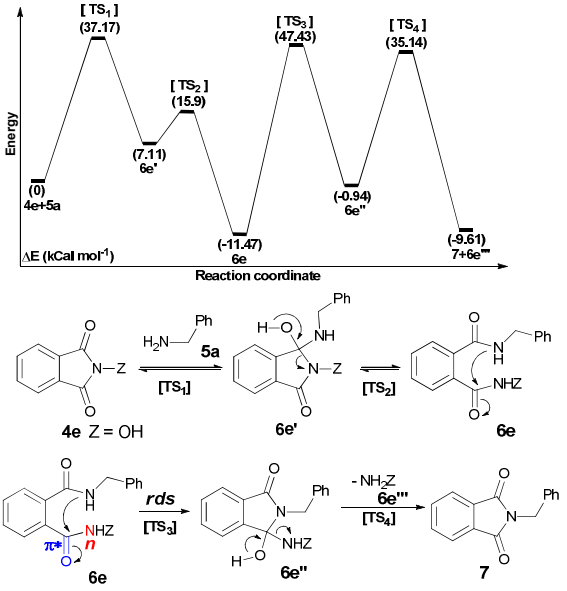
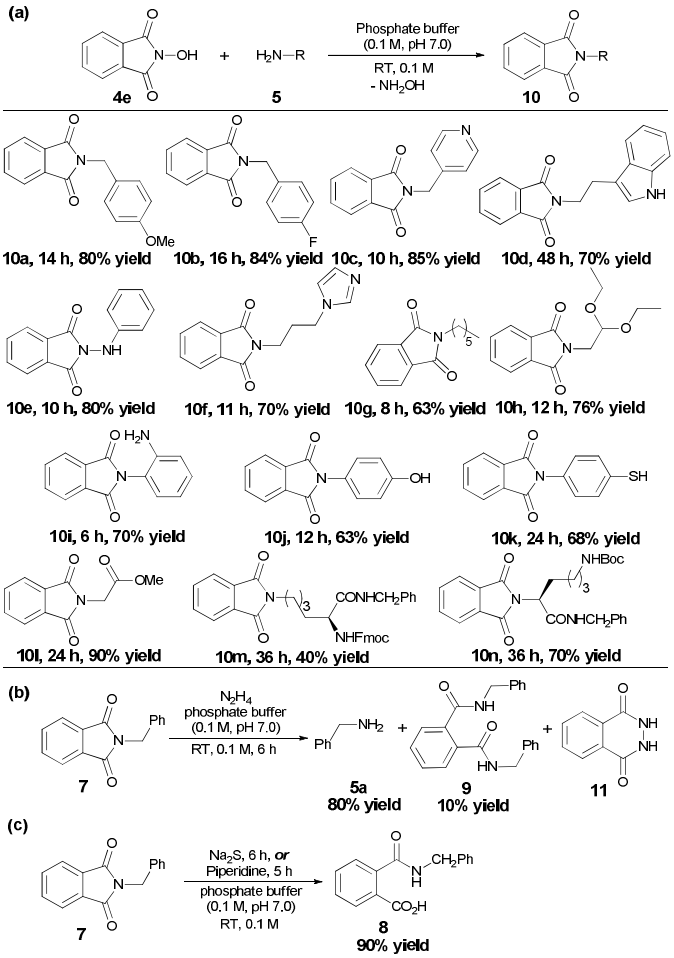


Figure 1. Potential energy curve for phthalimidation of **5a** with **4e**.



Scheme 3. (a) Phthalimidation of amines, (b) and (c) transformations of phthalimidoamine **7**.

The reaction with electronically diverse set of substrates underwent smoothly and the chemoselectivity was apparent with **10i**, **10j** and

10k. Aromatic amines are sparingly soluble in aqueous buffer but they dissolve with the progress of reaction. Degassing of solvent was found necessary for **10k** in improving the yields from 53% to 68%. The preferential reactivity of α-amine with respect to ε-amine in an amino acid was apparent with **10l**, **10m** and **10n**.

The phthalimidation protocol works smoothly in the pH range 6.0-7.8 to give product **7** in >99% conversion (Table S2, ESI). Efficiency of this transformation is not compromised at various scales ranging from 0.05 mmol to 50 mmol where product was isolated from 10 mg to 8.5 g (Table S3, ESI). Whereas benzyl amine reacts with NHP in 14 h to give high yields of the product, benzyl alcohol and benzyl thiol gave no conversions within 24 h. The phthalimidoamine product can be conveniently converted back to amine in presence of hydrazine (Scheme 3b).[14c] Treatment of the product with Na_2S or piperidine can catalyze the ring opening of phthalimidoamine followed by hydrolysis to render aromatic carboxylic acid (Scheme 3c).

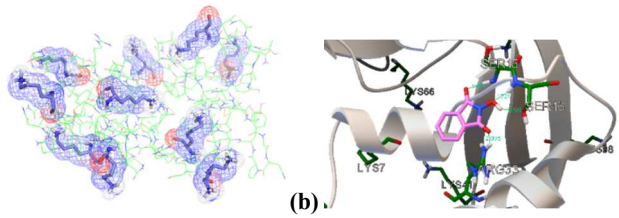
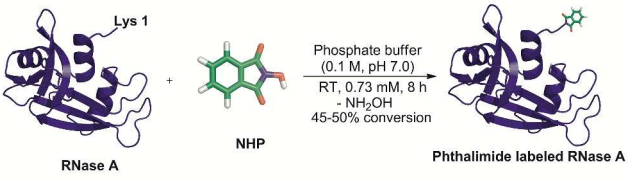


Figure 2. (a) NMR structure of RNase A[18] with highlighted ten lysine residues. (b) Docking investigations of NHP and RNase A.

We selected RNase A as a model protein for examining the efficiency of our methodology in protein labeling. The protein comprises of all the natural amino acids except tryptophan. It includes ten lysine residues where most of them are solvent exposed (Figure 2a). Five equivalents of NHP with RNase A resulted in mono-, bis- and tris-labeled products within 24 h (for details, see ESI). An explicit chemoselectivity for primary amine residue was evident with no apparent labeling of competing functionalities. In addition, the difference in reactivity of multiple copies of primary amine was substantiated. The relative pKa values of α-amine and ε-amine suggest that the prior would be more nucleophilic in the given conditions. We argued that selectivity between ten copies of ε-amine can result due to binding of NHP to an energetically favorable site where labeling would be dependent on the binding specificity. According to this hypothesis, K41 in RNase A active site will be labeled preferentially, whereas K66 and K7 can offer the second choice (docking studies, Figure 2b and ESI). The alternative rationale would allow the solvent accessible amine (K1, K66, K91)[5] to be labeled under kinetic control. The mass analysis of digested protein suggests that the three modified residues are K1 or K7, K37 or K41 and K91. This result indicates that all the three factors discussed above might be at interplay. Realizing the potential of methodology, we examined the reaction with equimolar NHP. Interestingly, only monolabeled product (45-50% conversion) was formed in 8 h (Scheme 4).[19]



Scheme 4. Chemoselective and site-selective phthalimidation of RNase A.

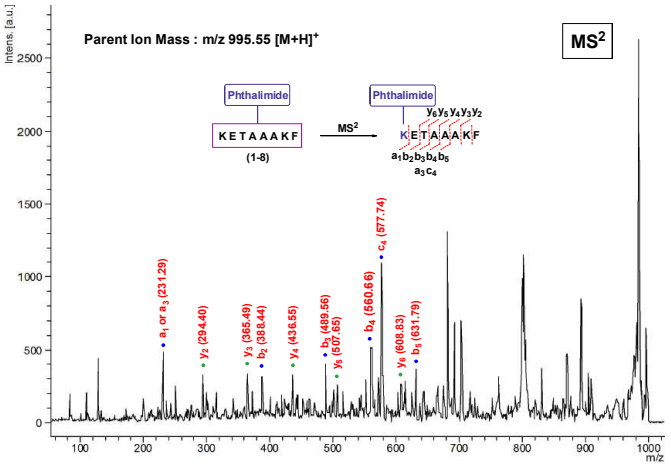


Figure 3. MS^2 spectra of phthalimide labeled peptide fragment (K1-F8) of RNase A.

In-solution digestion of modified RNase A by chymotrypsin resulted in peptide fragments where mass change of single fragment (K1-F8: KETAAAKF) can be observed from 865 Da to 995 Da. Presence of two Lys residues necessitated MS^2 experiment that confirmed K1 modification by presence of b2, b3, b4, b5, y2, y3, y4, y5 among other fragments (Figure 3). Formation of heterogeneously labeled protein was ruled out with the support of MS and MS^2 data. We believe that α-amine is labeled in K1 as its inherent nucleophilicity is higher than ε-amine at pH 7.0. The unique reactivity of α-amine in the protein backbone renders the site-selectivity.[4,20] The proposal is further supported by selective labeling of α-amine in a tetrapeptide GGGK-NH_2 when it was subjected to phthalimidation and characterized by MS and MS^2 (fragments b2, b3, y1, y2, z1, z2; Figure S11, ESI). The results were further substantiated by experiments with HGGK-NH_2, NGGK-NH_2, MGK-NH_2, and SFK-NH_2.[22]

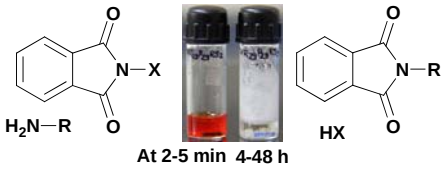
In conclusion, we have developed a phthalimidation protocol in aqueous buffer that provides good yields with range of substrates. The control over reactivity of the amphoteric intermediate was found to be the key in regulating chemoselectivity. Analytically pure products were isolated without column chromatography, in general. The reaction can be scaled up to multigram scale without compromising the reactivity and selectivity. The mild reaction conditions allowed us to avoid side reactions involving functionalities of other amino acid residues. RNase A gave 45-50% conversion to homogeneously monolabeled conjugate with one equivalent of reagent within 8 h. The transformation does not require large excess of the reagent.[21] Bimolecular protein conjugation is generally plagued by hydrolysis that follows pseudo first order kinetics. We have demonstrated that it can be evaded by shifting the rate-determining step to an intramolecular process. Development of N-hydroxyphthalimide derivatives that can offer fluorescent probes (Figures S16 and S17, Scheme S2, ESI) and precursor to the click chemistry is currently under investigation in our laboratory.

R.S., S.R.A., and P.N.J. are recipient of research fellowship from CSIR, India. V.R. is a Ramanujan Fellow (SERB, India). The authors are thankful to SERB, DAE and DBT, India for research grants.

Notes and references

[a] Department of Chemistry, Indian Institution of Science Education and Research (IISER) Bhopal, Indore By-pass Road, Bhauri, Bhopal, MP, India 462066. E-mail: vrai@iiserb.ac.in
† These authors contributed equally to this work.
Electronic Supplementary Information (ESI) available: Methods, procedures, additional results and discussion, characterization data, computation details. See DOI: 10.1039/b000000x/

1 (a) Y. X. Chen, G. Triola and H. Waldmann, *Acc. Chem. Res*., 2011, **44**, 762-773. (b) M. Chilamari and V. Rai, *Indian J. Chem. Sec A.,* 2013 **52A**, 992-1003.
2 J. E. Reardon, R. C. Crouch and L. St John Williams, *J. Biol. Chem*., 1994, **69**, 15999-16008.
3 O. Dmitrenko, C. Thorpe and R. D. Bach, *J. Org. Chem*., 2007, **72**, 8298-8307.
4 A. C. Obermeyer, J. B. Jarman and M. B. Francis, *J. Am. Chem. Soc*., 2014, **136**, 9572-9579.
5 For kinetic labeling of protein by acylation, see: X. Chen, K. Muthoosamy, A. Pfisterer, B. Neumann and T. Weil *Bioconjugate Chem.* 2012, **23**, 500-508.
6 (a) L. D. Lavis and R. T. Raines, *ACS Chem. Biol*., 2008, **3**, 142-155. (b) M. Pellecchia, D. S. Sem and K. Wuthrich, *Nat. Rev. Drug Discov*., 2002, **1**, 211-219.
7 B. Rup and D. O'Hara, *AAPS J*., 2007, **9**, E148-E155.
8 C. A. Burtis and E. R. Ashwood, *Tietz Textbook of Clinical Chemistry*, 2nd Ed., (Ed.: W. B. Saunders), Philadelphia, PA, 1994.
9 P. Jonkheijm, D. Weinrich, H. Schröder, C. M. Niemeyer and H. Waldmann, *Angew. Chem. Int. Ed*., 2008, **47**, 9618-9647.
10 A. H. Dent and M. Aslam, *Bioconjugation*: *Protein Coupling Techniques for the Biomedical Sciences*, Macmillan Reference Ltd., London, 1998.
11 R. L. Lundblad, *Chemical Reagents for Protein Modification*, 3rd Ed., CRC Press, Boca Raton, FL, 2005.
12 G. T. Hermanson, *Bioconjugate Techniques*, 3rd Ed., Academic Press, San Diego, CA, 2013.
13 J. C. Jewett and C. R. Bertozzi, *Chem. Soc. Rev*., 2010, **39**, 1272-1279.
14 For reports on phthalimidation protocols, see (a) S. N. Rao, D. C. Mohan and S. Adimurthy, *Org. Lett.*, 2013, **15**, 1496-1499; (b) M. E. Vázquez, J. B. Blanco and B. Imperiali, *J. Am. Chem. Soc.*, 2005, **127**, 1300-1306; (c) M. Goodman, *Synthesis of Peptides and Peptidomimetics*, E22a, Houben-Weyl, Thieme, Stuttgart, 2004.
15 X. Guo, F. S. Kim, S. A. Jenekhe and M. D. Watson, *J. Am. Chem. Soc*., 2009, **131**, 7206-7207.
16 G. Loving and B. Imperiali, *J. Am. Chem. Soc*., 2008, **130**, 13630-13638.
17 (a) G. Loving and B. Imperiali, *Bioconjugate Chem*., 2009, **20**, 2133-2141. (b) J. Riedl, R. Pohl, N. P. Ernsting, P. Orsag, M. Fojta and M Hocek, *Chem. Sci.*, 2012, **3**, 2797-2806.
18 PDB ID: 2AAS. J. Santoro, C. Gonzalez, M. Bruix, J. L. Neira, J. L. Nieto, J. Herranz and M. Rico, *J. Mol. Biol*., 1993, **229**, 722-734.
19 Conversions were calculated by spectroscopic determination of unreacted NHP. The extraction of NHP was optimized for <5% error. For details, see supplementary information.
20 For site-selective N-terminal modification of proteins, also see: (a) A. O.-Y. Chan, C.-M. Ho, H.-C. Chong, Y.-C. Leung, J.-S. Huang, M.-K. Wong and C.-M. Che, *J. Am. Chem. Soc.*, 2012, **134**, 2589-2598. (b) H. B. F. Dixon, *J. Protein Chem.*, 1984, **3**, 99-108.
21 Large excess of reagents (10-1000 equivalents) are typically required for bioconjugate transformations. See reference 12 for details.
22 For experimental details and data, see ESI (Figures S12-15).



- No catalyst or co-reagent
- Ambient Temp. / Aq. buffer (pH 7.0)
- Reaction time: 4-48 h
- Convenient purification / Scalable