ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

A highly efficient oxidative condensation reaction for selective protein conjugation†

Cite this: *Chem. Commun.*, 2014, **50**, 7469

Received 28th February 2014,
Accepted 22nd May 2014

DOI: 10.1039/c4cc01551g

www.rsc.org/chemcomm

Ao Ji,‡ Wei Ren‡ and Hui-wang Ai*

We hereby report a mild and efficient coupling reaction between alkyl aldehydes and aryl diamines. In the presence of a Cu^{2+} or a Zn^{2+} ion, oxygen (O_2) in air is able to promote the oxidative condensation of the two readily preparable functional groups, forming stable benzimidazole linkages in neutral aqueous solution at room temperature (RT). We demonstrated that the reaction could be utilized to label a T4 lysozyme protein containing a chemically installed aryl diamine group with a fluorescent aldehyde dye molecule at 37 °C.

In the past few decades, chemical coupling reactions that are compatible with common biomolecules have attracted intensive attention.[1] Examples (Table S1, ESI†) include azide–alkyne cycloaddition,[2] thiol–ene coupling reactions,[3] the Staudinger–Bertozzi ligation,[4] a photoclick tetrazole–alkene cycloaddition,[5] the cyanobenzothiazole condensation,[6] and a few other Diels–Alder cycloaddition reactions including the ones between tetrazines and ring-strained alkenes or alkynes.[7] These reactions have now been widely utilized to assemble organic molecules and molecular libraries,[8] label biomolecules with fluorescent dyes and other biophysical probes,[9] and conjugate biomolecules with polyethylene glycols (PEGs), radioactive isotopes, toxins or drugs for medical applications.[10] Despite the progress, a further expanded set of bioconjugation reactions is still in high demand to match diverse applications.[1,11] Furthermore, new expansion would provide additional, and sometimes better, choices when multiple mutually compatible reactions are employed to modify one or several biomolecules.[12]

Aldehydes may interact with biological nucleophiles such as amines and thiols, but the equilibrium in water favors the reverse hydrolytic reaction.[13] Due to its small size and the availability of selective reactions, aldehyde has been utilized as a reactive chemical handle to label biomolecules *in vitro* and in living cells. In most existing examples, molecules containing an alkoxyamine, a hydrazide or a hydrazine functional group were utilized to react with aldehyde.[9a,14] However, these conjugation reactions often suffer from drawbacks such as the favoring of acidic conditions and incomplete conversion due to the reversible feature.[15] To address the problems, anilines have been developed as nucleophilic catalysts to assist bioconjugation at neutral pH,[15b,16] but the general effectiveness of these catalysts remains to be investigated. Under some circumstances, aniline catalysts were reported to be neutral or even detrimental to reaction kinetics and equilibrium.[14a,17] Recently, a modified Pictet–Spengler reaction has been developed:[15a,18] small molecules were conjugated to aldehyde-containing proteins using the modified reaction to afford stable irreversible linkages, but a slightly acidic pH range was still required. In this context, we set to develop additional aldehyde coupling reactions that can be robustly operated in neutral aqueous solution. Herein we describe our recent effort in developing an oxidative condensation reaction between alkyl aldehydes and aryl diamines, and employing this highly efficient and mild coupling reaction to site-specifically label a model protein at the physiological pH and 37 °C.

The direct condensation of aldehydes and aryl diamines is a popular route for synthesizing benzimidazoles, a group of heterocyclic aromatic compounds known for their antiviral, antimitotic and anthelmintic properties.[19] The aryl diamine functional group is unreactive to all natural functional groups in proteins, which is a favorable property for applications in bioconjugation. However, existing condensation methods often require conditions (*e.g.* organic solvents, strong acids, elevated temperature, or microwave) that are too harsh for biomolecular labeling.[20] In one example, boric acid was reported to catalyze the condensation between *o*-phenylenediamine and aromatic aldehydes in water at RT.[21] However, we failed to obtain the reported benzimidazole products. Recently, Jiao and coworkers described an elegant 12 hour catalyst-free condensation of aryl diamines and cyclopropyl aldehydes in toluene,[19b] a solvent inappropriate for bioconjugation. To identify reaction conditions suitable for biomolecular labeling, we examined various combinations of aldehydes, diamines and common Lewis acid catalysts in aqueous phosphate buffer (50 mM, pH 7.4). We were fortunate to find that, in the presence of oxygen (O_2) in air and

Department of Chemistry, University of California Riverside, 501 Big Springs Road, Riverside, CA 92521, USA. E-mail: huiwang.ai@ucr.edu

† Electronic supplementary information (ESI) available: General methods for organic synthesis and protein labeling, SDS-PAGE analysis and spectroscopic characterization. See DOI: 10.1039/c4cc01551g

‡ These two authors contributed equally to this work.

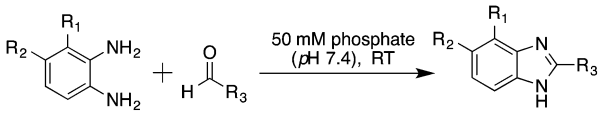
a catalytic amount of Cu^{2+} or Zn^{2+} metal ion, aryl diamines and alkyl aldehydes can react quickly to form corresponding benzimidazoles in good yields at RT.

We further explored the substrate scope of this condensation reaction. Alkyl aldehyde seemed to be required. The reaction between benzaldehyde and *o*-phenylenediamine gave no benzimidazole product (entry 1, Table 1). In comparison, benzimidazoles could be synthesized from various aryl diamines and butyraldehyde (entries 2–5). In general, electron-donating groups (*e.g.* $-OCH_3$ and $-CH_3$) substituted on the diamine aryl ring can accelerate the reaction and enhance the yields, while electron-withdrawing groups (*e.g.* $-NO_2$ and –COOMe) disfavor the formation of benzimidazoles. The reaction between 4-methoxybenzene-1,2-diamine and *o*-phenylenediamine completed nearly-quantitatively in 30 min (entry 4), while 4-nitrobenzene-1,2-diamine under the same reaction condition resulted in no significant amount of benzimidazole over 10 hours (entry 6). The reaction with cyclopropanecarboxaldehyde (entry 7) progressed equally well as the one with butyraldehyde (entry 3). It is also worth noting that, these condensation reactions (entries 2–5 and 7) can be performed in phosphate buffer, as well as in neat water and organic solvents such as tetrahydrofuran (THF), as long as Cu^{2+} or Zn^{2+} is supplemented. Both metal ions and O_2 in air are essential for the excellent reaction yields to form benzimidazoles (entry 2).

We next investigated the kinetics of the metal ion-catalyzed reaction. A mixture of 4-methoxybenzene-1,2-diamine and butyraldehyde was stirred in neutral phosphate buffer containing 10 μM $CuSO_4$ in open air at RT. In a 50 min time period, portions of the reaction mixture were removed and processed immediately for NMR analysis. A characteristic change in proton chemical shift of the methoxy group substituted in the aryl ring was observed (Fig. 1a). By integrating the areas of the corresponding peaks, we determined the molar concentrations of the product and the reactants at each time point. A second-order kinetics model was utilized to fit the data (Fig. 1b), and the rate constant of the coupling reaction was determined to be 1.61 M^{-1} s^{-1}. The rate of the reaction is comparable to many commonly used bioconjugation reactions (Table S1, ESI†).[12b,22]

The two catalysts for the coupling reaction, Cu^{2+} and Zn^{2+}, are essential elements for life and have reasonable compatibility with biomolecules.[1,23] We next explored the use of the oxidative condensation reaction to site-specifically label proteins. We acknowledge that a variety of elegant chemical, enzymatic and genetic methods may possibly be adapted to introduce alkyl aldehyde or aryl diamine functional groups into proteins.[9a,24] In our model experiment, we used a cysteine–maleimide Michael reaction to chemically install an aryl diamine functional group to a T4 lysozyme protein. Briefly, T4 lysozyme containing a single cysteine at the surface residue 68 (T4L-N68C) was recombinantly expressed in *E. coli*. A bifunctional linker molecule containing an aromatic diamine and a cysteine-reactive maleimide was chemically synthesized (Scheme S1, ESI†) and utilized to react with T4L-N68C. The successful conversion (Fig. 2a and b) was confirmed using

Table 1 Oxidative coupling of aldehydes and aryl diamines in the presence of a Cu^{2+} or a Zn^{2+} catalyst (1 mol%)



1 2 3a-g

Entry		R_1	R_3	Molar ratio[a] (diamine : aldehyde)	Reaction time	Product	Catalyst	Isolated yield[b] (%)
1	H	H	Ph	1 : 2	5 h	**3a**	$CuSO_4$	Air: 0
							$ZnCl_2$	Air: 0
2	H	H	*n*-Propyl	1 : 2	1.5 h	**3b**	$CuSO_4$	Air: 84
							$ZnCl_2$	Air: 66
							No catalyst	Air: 15
							$CuSO_4$	N_2: <5[c]
3	Me	H	*n*-Propyl	1 : 2	1 h	**3c**	$CuSO_4$	Air: 84
							$ZnCl_2$	Air: 63
4	H	Ome	*n*-Propyl	1 : 2	30 min	**3d**	$CuSO_4$	Air: 92
							$ZnCl_2$	Air: 94
5	H	COOMe	*n*-Propyl	1 : 5	4 h	**3e**	$CuSO_4$	Air: 63
							$ZnCl_2$	Air: 55
6	H	NO_2	*n*-Propyl	1 : 2	10 h	**3f**	$CuSO_4$	Air: 0
							$ZnCl_2$	Air: 0
7	Me	H	Cyclopropyl	1 : 2	1 h	**3g**	$CuSO_4$	Air: 81
							$ZnCl_2$	Air: 67

[a] Additional volatile aldehydes were used to maximize diamine conversion. [b] Isolated yields were calculated from the moles of the benzimidazole products after column chromatography and the initial moles of aryl diamines. [c] Residual benzimidazole product may form during workup procedures performed in air.

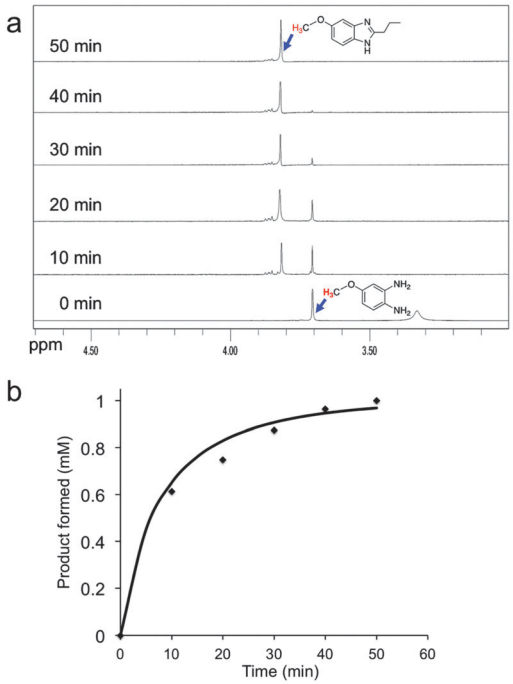


Fig. 1 (a) ^{1}H NMR time course showing the conversion of 4-methoxybenzene-1,2-diamine to its corresponding benzimidazole. (b) Reaction kinetics of the formation of the benzimidazole product over time.

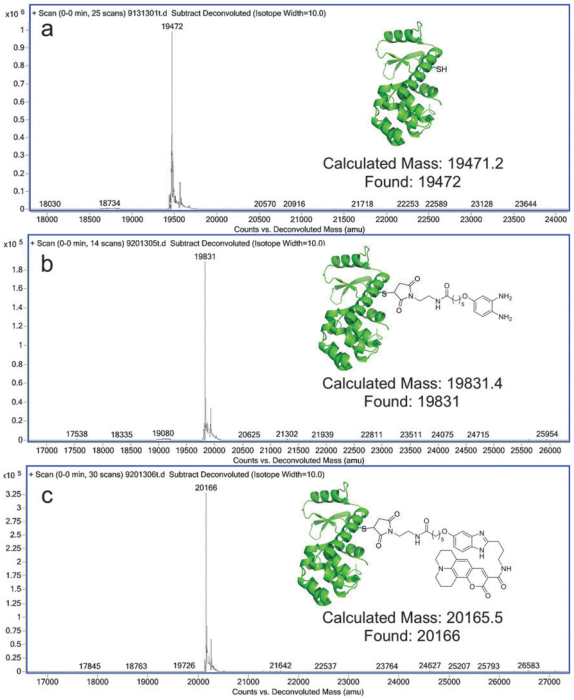


Fig. 2 Direct infusion ESI-TOF MS characterization of T4 lysozyme proteins ((a) T4L-N68C; (b) T4L containing an installed aryl diamine functional group; (c) T4L labeled with an aldehyde Coumarin 343 dye in the presence of 2.5 μM $CuSO_4$).

electrospray ionization-time of flight mass spectrometry (ESI-TOF MS). After obtaining the T4L protein with an installed aryl diamine group, we tested its reaction with an aldehyde-containing Coumarin 343 dye (Schemes S2 & S3, ESI†). The protein (50 μM) was incubated with the dye (200 μM) in phosphate buffer (50 mM, pH 7.4) supplemented with $CuSO_4$ (2.5 μM). The mixture was left in open air at 37 °C. The reaction was monitored with ESI-TOF MS in a direct infusion mode. It showed that the protein was quantitatively converted to a fluorescently labeled product in 2 hours (Fig. 2c). In the control experiment when T4L-N68C and the aldehyde were directly mixed, no reaction was observed (Fig. S1a, ESI†). The proteins were also analyzed on a SDS-PAGE gel (Fig. S2a, ESI†). A fluorescent band was seen for the Coumarin 343-labeled protein. After removing the excess dye, the fluorescence of the protein was characterized (Fig. S2b, ESI†). The excitation and emission spectra matched the cyan fluorescence of Coumarin 343. Efficient labeling was also observed in the presence of $ZnCl_2$, but not in the absence of metal ions (Fig. S1 and S2, ESI†). All these results collectively support that the oxidative condensation reaction is a potent and selective method that can be utilized for site-specific protein conjugation.

In summary, we have identified a mild condition to couple aryl diamines and alkyl aldehyde in neutral aqueous solution. The resulting benzimidazole derivatives are a family of important bioactive compounds, whose preparation is being actively pursued by many researchers.[19b,25] The one-step transformation reported here is a very effective way to prepare these compounds. The coupling reaction has favorable kinetics, excellent conversion and good biocompatibility. We successfully used it to label a protein containing a pre-installed aryl diamine functional group. Since the two coupling functional groups are readily accessible and small in structure, broad applications of this novel bioconjugation reaction can be expected. We are currently developing strategies to genetically encode aldehydes and diamines in living cells, which are expected to further ease the use of the condensation reaction for selective protein modification. Because the metal ion concentration used here is considerably lower than that in previous studies which reduced Cu^{2+} for live-cell alkyne–azide click reactions,[1,14b,26] an integration of this condensation reaction with strategies to introduce the functional groups into biomolecules may permit site-specific labeling of biomolecules on the surface of or inside living cells.

We gratefully acknowledge support from the University of California Riverside. We thank the UCR NMR Facility and the High Resolution Mass Spectrometry Facility for analysis of small molecules and proteins.

Notes and references

1 E. M. Sletten and C. R. Bertozzi, *Angew. Chem., Int. Ed.*, 2009, **48**, 6974–6998.
2 (*a*) C. W. Tornoe, C. Christensen and M. Meldal, *J. Org. Chem.*, 2002, **67**, 3057–3064; (*b*) V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2002, **41**, 2596–2599; (*c*) 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.
3 C. E. Hoyle and C. N. Bowman, *Angew. Chem., Int. Ed.*, 2010, **49**, 1540–1573.
4 E. Saxon and C. R. Bertozzi, *Science*, 2000, **287**, 2007–2010.
5 R. K. Lim and Q. Lin, *Acc. Chem. Res.*, 2011, **44**, 828–839.
6 H. Ren, F. Xiao, K. Zhan, Y. P. Kim, H. Xie, Z. Xia and J. Rao, *Angew. Chem., Int. Ed.*, 2009, **48**, 9658–9662.

7 (*a*) N. K. Devaraj and R. Weissleder, *Acc. Chem. Res.*, 2011, **44**, 816–827; (*b*) K. Lang, L. Davis, J. Torres-Kolbus, C. Chou, A. Deiters and J. W. Chin, *Nat. Chem.*, 2012, **4**, 298–304; (*c*) 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.
8 (*a*) H. C. Kolb and K. B. Sharpless, *Drug Discovery Today*, 2003, **8**, 1128–1137; (*b*) S. L. Ng, P. Y. Yang, K. Y. Chen, R. Srinivasan and S. Q. Yao, *Org. Biomol. Chem.*, 2008, **6**, 844–847; (*c*) V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano and M. F. Greaney, *Nat. Chem.*, 2010, **2**, 490–497.
9 (*a*) P. Wu, W. Shui, B. L. Carlson, N. Hu, D. Rabuka, J. Lee and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 3000–3005; (*b*) J. C. Jackson, J. T. Hammill and R. A. Mehl, *J. Am. Chem. Soc.*, 2007, **129**, 1160–1166; (*c*) V. Muralidharan and T. W. Muir, *Nat. Methods*, 2006, **3**, 429–438; (*d*) S. T. Laughlin and C. R. Bertozzi, *Nat. Protocols*, 2007, **2**, 2930–2944; (*e*) H. W. Ai, *Anal. Bioanal. Chem.*, 2012, **403**, 2089–2102.
10 (*a*) H. Cho, T. Daniel, Y. J. Buechler, D. C. Litzinger, Z. Maio, A. M. Putnam, V. S. Kraynov, B. C. Sim, S. Bussell, T. Javahishvili, S. Kaphle, G. Viramontes, M. Ong, S. Chu, G. C. Becky, R. Lieu, N. Knudsen, P. Castiglioni, T. C. Norman, D. W. Axelrod, A. R. Hoffman, P. G. Schultz, R. D. DiMarchi and B. E. Kimmel, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 9060–9065; (*b*) C. H. Kim, J. Y. Axup, A. Dubrovska, S. A. Kazane, B. A. Hutchins, E. D. Wold, V. V. Smider and P. G. Schultz, *J. Am. Chem. Soc.*, 2012, **134**, 9918–9921; (*c*) K. Nwe and M. W. Brechbiel, *Cancer Biother. Radiopharm.*, 2009, **24**, 289–302.
11 C. R. Bertozzi, *Acc. Chem. Res.*, 2011, **44**, 651–653.
12 (*a*) M. D. Best, *Biochemistry*, 2009, **48**, 6571–6584; (*b*) R. K. Lim and Q. Lin, *Chem. Commun.*, 2010, **46**, 1589–1600.
13 W. P. Jencks, *J. Am. Chem. Soc.*, 1959, **81**, 475–481.
14 (*a*) X. Shi, Y. Jung, L. J. Lin, C. Liu, C. Wu, I. K. Cann and T. Ha, *Nat. Methods*, 2012, **9**, 499–503; (*b*) J. D. Cohen, P. Zou and A. Y. Ting, *ChemBioChem*, 2012, **13**, 888–894; (*c*) L. Yi, H. Sun, Y. W. Wu, G. Triola, H. Waldmann and R. S. Goody, *Angew. Chem., Int. Ed.*, 2010, **49**, 9417–9421.
15 (*a*) P. Agarwal, R. Kudirka, A. E. Albers, R. M. Barfield, G. W. de Hart, P. M. Drake, L. C. Jones and D. Rabuka, *Bioconjugate Chem.*, 2013, **24**, 846–851; (*b*) A. Dirksen and P. E. Dawson, *Bioconjugate Chem.*, 2008, **19**, 2543–2548.
16 (*a*) A. Dirksen, T. M. Hackeng and P. E. Dawson, *Angew. Chem., Int. Ed.*, 2006, **45**, 7581–7584; (*b*) P. Crisalli and E. T. Kool, *J. Org. Chem.*, 2013, **78**, 1184–1189; (*c*) M. Rashidian, M. M. Mahmoodi, R. Shah, J. K. Dozier, C. R. Wagner and M. D. Distefano, *Bioconjugate Chem.*, 2013, **24**, 333–342; (*d*) P. Crisalli and E. T. Kool, *Org. Lett.*, 2013, **15**, 1646–1649; (*e*) Y. Zeng, T. N. Ramya, A. Dirksen, P. E. Dawson and J. C. Paulson, *Nat. Methods*, 2009, **6**, 207–209.
17 (*a*) Y. Taga, M. Kusubata, K. Ogawa-Goto and S. Hattori, *Mol. Cell. Proteomics*, 2012, **11**, M111.010397; (*b*) J. E. Hudak, R. M. Barfield, G. W. de Hart, P. Grob, E. Nogales, C. R. Bertozzi and D. Rabuka, *Angew. Chem., Int. Ed.*, 2012, **51**, 4161–4165.
18 P. Agarwal, J. van der Weijden, E. M. Sletten, D. Rabuka and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 46–51.
19 (*a*) M. Alamgir, D. C. Black and N. Kumar, in *Bioactive Heterocycles III*, ed. M. Khan, Springer, Berlin, Heidelberg, 2007, ch. 88, vol. 9, pp. 87–118; (*b*) C. Zhang, L. Zhang and N. Jiao, *Green Chem.*, 2012, **14**, 3273–3276; (*c*) D. A. Horton, G. T. Bourne and M. L. Smythe, *Chem. Rev.*, 2003, **103**, 893–930.
20 (*a*) L. M. Dudd, E. Venardou, E. Garcia-Verdugo, P. Licence, A. J. Blake, C. Wilson and M. Poliakoff, *Green Chem.*, 2003, **5**, 187–192; (*b*) J.-P. Wan, S.-F. Gan, J.-M. Wu and Y. Pan, *Green Chem.*, 2009, **11**, 1633–1637; (*c*) C. Mukhopadhyay and P. K. Tapaswi, *Tetrahedron Lett.*, 2008, **49**, 6237–6240; (*d*) K. Bahrami, M. M. Khodaei and F. Naali, *J. Org. Chem.*, 2008, **73**, 6835–6837; (*e*) Y.-X. Chen, L.-F. Qian, W. Zhang and B. Han, *Angew. Chem., Int. Ed.*, 2008, **47**, 9330–9333.
21 Z. Karimi-Jaberi and M. Amiri, *E-J. Chem.*, 2012, **9**, 167–170.
22 S. I. Presolski, V. Hong, S.-H. Cho and M. G. Finn, *J. Am. Chem. Soc.*, 2010, **132**, 14570–14576.
23 K. M. Dean, Y. Qin and A. E. Palmer, *Biochim. Biophys. Acta*, 2012, **1823**, 1406–1415.
24 (*a*) M. Rashidian, J. K. Dozier and M. D. Distefano, *Bioconjugate Chem.*, 2013, **24**, 1277–1294; (*b*) M. Rashidian, J. M. Song, R. E. Pricer and M. D. Distefano, *J. Am. Chem. Soc.*, 2012, **134**, 8455–8467; (*c*) M. Rashidian, J. K. Dozier, S. Lenevich and M. D. Distefano, *Chem. Commun.*, 2010, **46**, 8998–9000.
25 (*a*) S.-K. Xiang, W. Tan, D.-X. Zhang, X.-L. Tian, C. Feng, B.-Q. Wang, K.-Q. Zhao, P. Hu and H. Yang, *Org. Biomol. Chem.*, 2013, **11**, 7271–7275; (*b*) T. Xiao, S. Xiong, Y. Xie, X. Dong and L. Zhou, *RSC Adv.*, 2013, **3**, 15592–15595.
26 C. Besanceney-Webler, H. Jiang, T. Zheng, L. Feng, D. Soriano del Amo, W. Wang, L. M. Klivansky, F. L. Marlow, Y. Liu and P. Wu, *Angew. Chem., Int. Ed.*, 2011, **50**, 8051–8056.

