nature chemistry

ARTICLES

PUBLISHED ONLINE: 22 JUNE 2015 | DOI: 10.1038/NCHEM.2282

An oxazetidine amino acid for chemical protein synthesis by rapid, serine-forming ligations

Ivano Pusterla[1] and Jeffrey W. Bode[1,2]*

Amide-forming ligation reactions allow the chemical synthesis of proteins by the union of unprotected peptide segments, and enable the preparation of protein derivatives not accessible by expression or bioengineering approaches. The native chemical ligation (NCL) of thioesters and N-terminal cysteines is unquestionably the most successful approach, but is not ideal for all synthetic targets. Here we describe the synthesis of an Fmoc-protected oxazetidine amino acid for use in the α-ketoacid–hydroxylamine (KAHA) amide ligation. When incorporated at the N-terminus of a peptide segment, this four-membered cyclic hydroxylamine can be used for rapid serine-forming ligations with peptide α-ketoacids. This ligation operates at low concentration (100 μM–5 mM) and mild temperatures (20–25 °C). The utility of the reaction was demonstrated by the synthesis of S100A4, a 12 kDa calcium-binding protein not easily accessible by NCL or other amide-forming reactions due to its primary sequence and properties.

Access to proteins is essential for modern medicine and biological science. As a complement to recombinant protein expression, the chemical synthesis of proteins has emerged as a viable approach to some protein targets and offers advantages including homogeneity, the potential to incorporate any unnatural amino acid or post-translational modification, and the ability to prepare protein enantiomers[1–3]. The native chemical ligation (NCL) of thioesters and N-terminal cysteines—reported by Kent and co-workers over twenty years ago—has been a transformative advance in protein synthesis, but is not ideal for many synthetic targets[4]. To identify more general and complementary protein ligation reactions, numerous groups have pursued the development of novel methods and ligation partners[5].

Our own research efforts identified the union of α-ketoacids and hydroxylamines (KAHA ligation) as a promising new chemoselective protein ligation[6,7]. We have synthesized proteins including Pup, CspA, UFM1 and SUMO2/3 by KAHA ligations between C-terminal peptide α-ketoacids and segments bearing an N-terminal 5-oxaproline residue (5-Opr)[8]. This ligation has proven to be remarkably robust, but has limitations including the introduction of a non-canonical homoserine residue at the ligation site, the formation of esters as the primary ligation products[9], and a preference for relatively high concentrations (10–20 mM) and temperatures (50–60 °C). The success of the 5-Opr, a cyclic five-membered ring hydroxylamine, in the ligation appears to derive from ring strain; analogous acyclic hydroxylamines are unreactive. Based on this observation, we postulated that oxazetidine monomers—four-membered ring hydoxylamines—would address all of the remaining limitations of the KAHA ligation for protein synthesis by providing native serine or threonine residues, operating at lower concentrations and temperatures due to increased reactivity, and affording directly the amide products.

Here, we describe the first synthesis of an oxazetidine amino acid by a chemoselective and stereospecific rearrangement to form the four-membered ring hydroxylamine. When incorporated at the N-terminus of a peptide segment, this highly strained—but sufficiently stable—amino acid residue can be used for rapid serine-forming ligations with the KAHA amide-forming reaction. It operates at low concentrations (100 μM–5 mM) and mild temperatures (20–25 °C) with short reaction times. The utility of this amino acid was demonstrated in the synthesis of S100A4 (metastasin), a 12 kDa calcium-binding protein not easily accessible by NCL due to its primary sequence and properties.

Results

At the outset of our studies it was unclear whether the necessary 1,2-oxazetidines could be synthesized or if they would be stable despite the high inherent ring strain (25 kcal mol^{-1}, Fig. 1a)[10]—higher even than the well-studied three-membered ring oxaziridines (23 kcal mol^{-1}). Furthermore, no previous examples of unsubstituted (that is, N–H) oxazetidine had been reported[11] and only one example of a protected oxazetine (*N*-Boc) has appeared, although a single computational study suggested that such compounds might be stable[10].

Our early attempts at devising a synthetic entry to **1**, including intramolecular *N*-alkylation[12], intramolecular *C*-alkylation[13] and [2+2] cycloadditions, were complicated by the unique reactivity of the N–O bond in our system and the well-known difficulty of closing four-membered rings[14]. We reasoned that if a hydroxylamine-epoxide (Fig. 1b) underwent an intramolecular cyclization, it could potentially give the oxazetidine by a 4-*exo*-tet pathway rather than the disfavoured 5-*endo*-tet closure[15,16], as a similar strategy has proven successful for cyclobutane formation[17]. A carbamate-protected hydroxylamine **2** (Fig. 1b) was synthesized and subjected to cyclization conditions, which cleanly, but unfortunately, delivered **3** by 6-*exo*-tet ring closure of the Boc oxygen atom[18]. A similar cyclization, followed by ^{t}Bu migration to give **4**, occurred when **2** was treated with $BF_3{\cdot}OEt_2$ (Fig. 1b). We opted instead for *N*-benzyl protected substrate **5**, which cannot close to form a six-membered ring, and was prepared in good yield. During its preparation, we serendipitously found that it easily cyclized to the undesired isoxazoline **6**, presumably via the intermediacy of an alkyl chloride. Although we initially expected the desired cyclization to proceed under basic conditions, this substrate resisted even harsh treatments (for example, potassium hexamethyldisilylamide (KHMDS), 80 °C, in dimethyl sulfoxide (DMSO))[19]. Continued studies revealed a critical role for Lewis acid activation

[1]Department of Chemistry and Applied Biosciences, Laboratorium für Organische Chemie, Vladimir Prelog Weg 3, ETH Zürich, Zürich 8093, Switzerland. [2]Institute of Transformative Bio-Molecules (WPI-ITbM), Nagoya University, Chikusa, Nagoya 464-8602, Japan. *e-mail: bode@org.chem.ethz.ch

© 2015 Macmillan Publishers Limited. All rights reserved

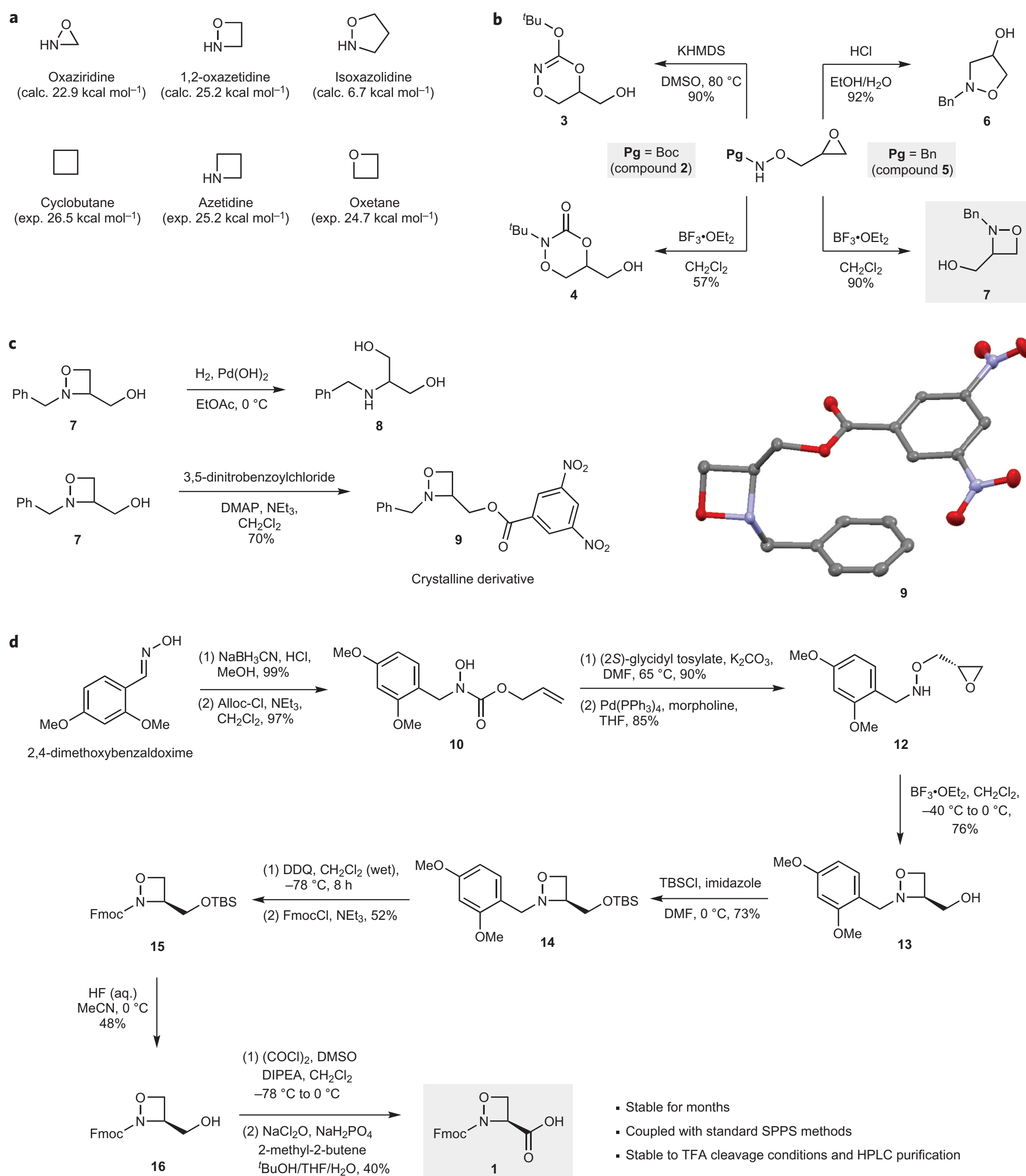
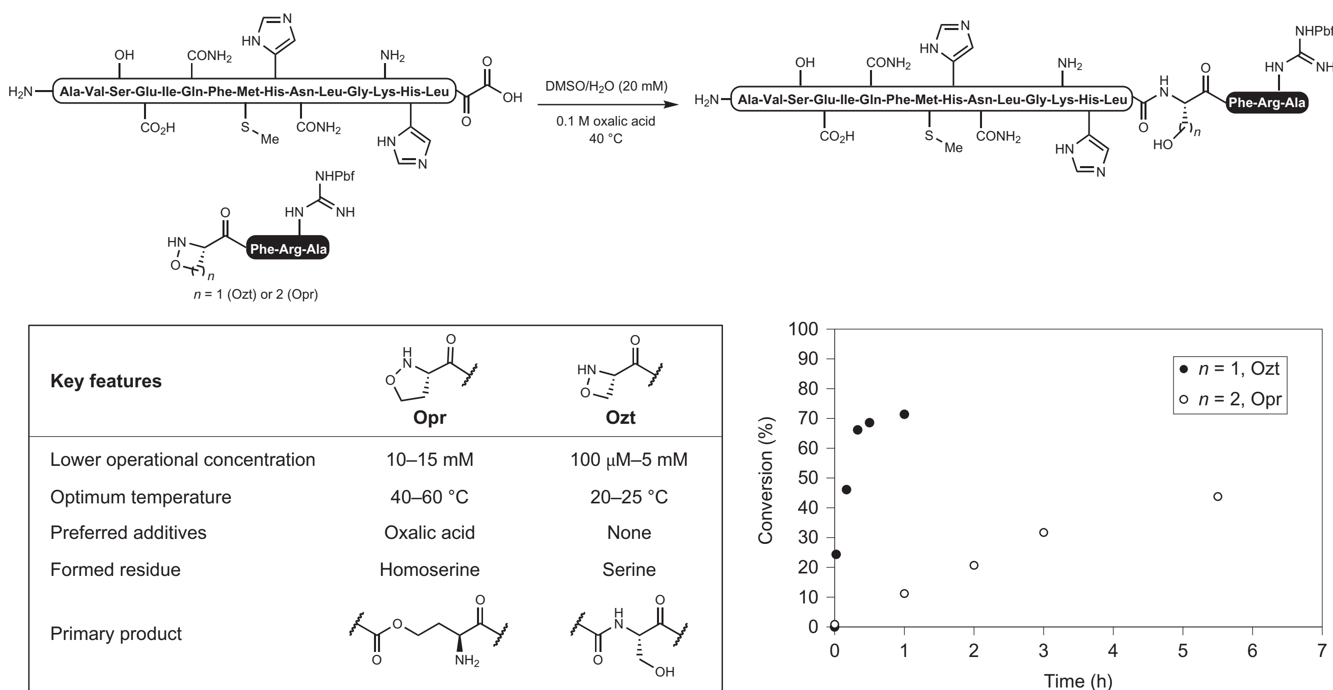


Figure 1 | Synthesis of enantioenriched Fmoc-oxazetidine amino acid 1. a, Ring strain in 1,2-oxazetidine and parent compounds. Calculated or experimental strain energies are shown in parentheses (see Supplementary Information for calculations). **b**, Cyclizations from the bifunctional epoxide-hydroxylamine. **c**, Characterization of the oxazetidine. **d**, Synthetic route to Fmoc-protected amino acid 1. DMAP, 4-dimethyaminopyridine; TBS, *tert*-butyldimethylsilyl; DDQ, 2,3-dichloro-5,6-dicyano-1,4-benzoquinone; Fmoc, 9-fluorenylmethyloxycarbonyl; DIPEA, diisopropylethylamine; TFA, trifluoroacetic acid.

of the epoxide, and treatment of 5 with $BF_3{\cdot}OEt_2$ in CH_2Cl_2 furnished the desired oxazetidine 7 in excellent yield. Unfortunately, while the *N*-Bn group could be easily removed from five-membered ring 6 without affecting the hydroxylamine, all attempts at *N*-Bn removal from 7 resulted in N–O bond cleavage, reflecting the much higher ring strain. This reduction confirmed the structure by comparison to known compound 8, an assignment further verified by X-ray analysis of ester derivative 9 (Fig. 1c). Unfortunately, these results exposed 7 as a dead-end for the synthesis of the *N*-Fmoc oxazetidine amino acid 1.

A literature survey of *N*-alkyl protecting groups that can be removed under neutral, non-reducing conditions provided little encouragement. Despite the fact that almost no examples of amine deprotections of oxidatively removable 2,4-dimethoxybenzyl were known[20], we selected this group in the hope of identifying conditions for its removal without destruction of the oxazetidine ring. After extensive optimization, we found that it could be removed by treatment with DDQ to give an *N*-unprotected oxazetidine that was immediately protected with FmocCl. On the basis of these results, we developed a scalable synthesis of enantioenriched monomer 1 by

© 2015 Macmillan Publishers Limited. All rights reserved



Key features	Opr	Ozt
Lower operational concentration	10–15 mM	100 μM–5 mM
Optimum temperature	40–60 °C	20–25 °C
Preferred additives	Oxalic acid	None
Formed residue	Homoserine	Serine
Primary product		

Figure 2 | A comparison between KAHA ligations with oxazetidine- and oxaproline-containing peptides. The plot shows relative conversion to the product at 20 mM, 40 °C and in the presence of oxalic acid (0.1 M), which are the optimal conditions for 5-Opr ligation.

alkylation of hydroxamic acid **10** with (2*S*)-(–)-glycidyl tosylate followed by cyclization with $BF_3{\cdot}OEt_2$. Remarkably—and fortunately—the cyclization proceeds via an S_N2 mechanism to give the (*S*)-configured oxazetidine product in 98% enantiomeric excess (e.e.). Protection of the primary alcohol, oxidative 2,4-dimethoxybenzyl removal and Fmoc protection provided **15**. Finally, tert-butyldimethylsilyl deprotection and a careful two-step oxidation of the alcohol gave the Fmoc-protected amino acid **1** (Fig. 1d).

With compound **1** in hand, we tested its stability and performance in KAHA ligations. We found that **1** is stable for months if kept at a low temperature (–20 °C) and could be coupled under standard solid phase peptide synthesis (SPPS) conditions using common coupling reagents. Interestingly, MALDI analysis of peptides containing the Fmoc-Ozt often showed the mass of a retro [2+2] cycloreversion, a fragmentation not observed by electron spray ionization analysis. We were pleased to find that the Fmoc protecting group could be removed with *N*,*N*-diethylamine without decomposition of the oxazetidine, although care in handling of the unprotected oxazetidine was needed and storage of the unprotected Ozt is not recommended. This is true only of the unprotected oxazetidine peptides; the Fmoc protected forms are completely stable to storage and other manipulation. When the deprotected oxazetidine-peptides were exposed to α-ketoacids in aqueous DMSO solutions, they underwent rapid and chemoselective ligation. Even at low concentrations (100 μM–1 mM) and at room temperature the reaction was complete within minutes. NMR and HPLC studies of a model peptide showed that the primary product of the oxazetidine ligation is the amide—no trace of the ester was detected—and the ligation furnished the product without detectable epimerization. A screening of the ligation conditions showed that the reaction worked well in aqueous *N*-methyl pyrrolidine (NMP) or DMSO with no additional additives. Under these conditions, the reaction proceeded to completion at room temperature within 30 min at 1 mM (Fig. 2). Competition experiments between peptides bearing 5-Opr and Ozt under standard KAHA conditions (20 mM, 40 °C) exclusively gave the Ozt-ligated product, suggesting that the Ozt ligation is at least 100 times faster. Treatment of the data shown in Fig. 2 gives an approximate second-order rate constant for Ozt of ~0.1 $M^{-1}\ s^{-1}$, while for 5-Opr it is ~0.001 $M^{-1}\ s^{-1}$. For comparison, the second-order rate constant of NCL is ~0.3 $M^{-1}\ s^{-1}$.

We applied **1** to the synthesis of human S100A4 (also known as metastasin)[21,22], a 100-residue, calcium-binding protein of the S100 family[23]. S100A4 contains four cysteine residues in its primary sequence (Cys3, Cys76, Cys81, Cys86), but these are not in strategic positions for NCL and are critical to the folding and function of the protein. The required segments were prepared by SPPS according to the strategy showed in Fig. 3. Three mutations were into the synthetic protein (Thr39Ser, Asn68Gly, Ser80Hse). The Ser80Hse mutation arose from early attempts that were plagued by poor solubility of the intermediates. These difficulties were overcome by taking advantage of the ester-forming ligation of 5-Opr and keeping the assembled segments as the ester until folding into a more soluble form. The Asn68Gly was chosen due to large amounts of aspartimide formation from the neighbouring Asp during the synthesis of segment two and the fact that most other S100 proteins have Gly at this residue. The Thr39Ser mutation allowed us to both demonstrate the ability to carry the protected Ozt-residue through multiple ligation steps and conduct the final ligation under lower (1–5 mM) concentrations.

The first KAHA ligation between segment 2 α-ketoacid (**17**) bearing the Fmoc-protected Ozt and segment 3 5-oxaproline **18** furnished the ligated ester, which was kept in this form to improve handling and solubility[24,25]. Fmoc-deprotection revealed N-terminal oxazetidine **20**, which was ligated with two equivalents of segment 1 α-ketoacid **21**. As hoped, the reaction gave the desired product **22** at mild temperatures (22 °C) and at low concentrations (2–5 mM), which circumvented aggregation or solubility issues, and the ligation was complete within hours. The unprotected amines, carboxylic acids and thiols present in the peptide segments did not interfere with the ligation or attack the oxazetidine ring. The identity and purity of **22** were confirmed by high-resolution mass-spectrometry and HPLC. Treatment of **22** with basic carbonate buffer resulted in Fmoc-deprotection

© 2015 Macmillan Publishers Limited. All rights reserved

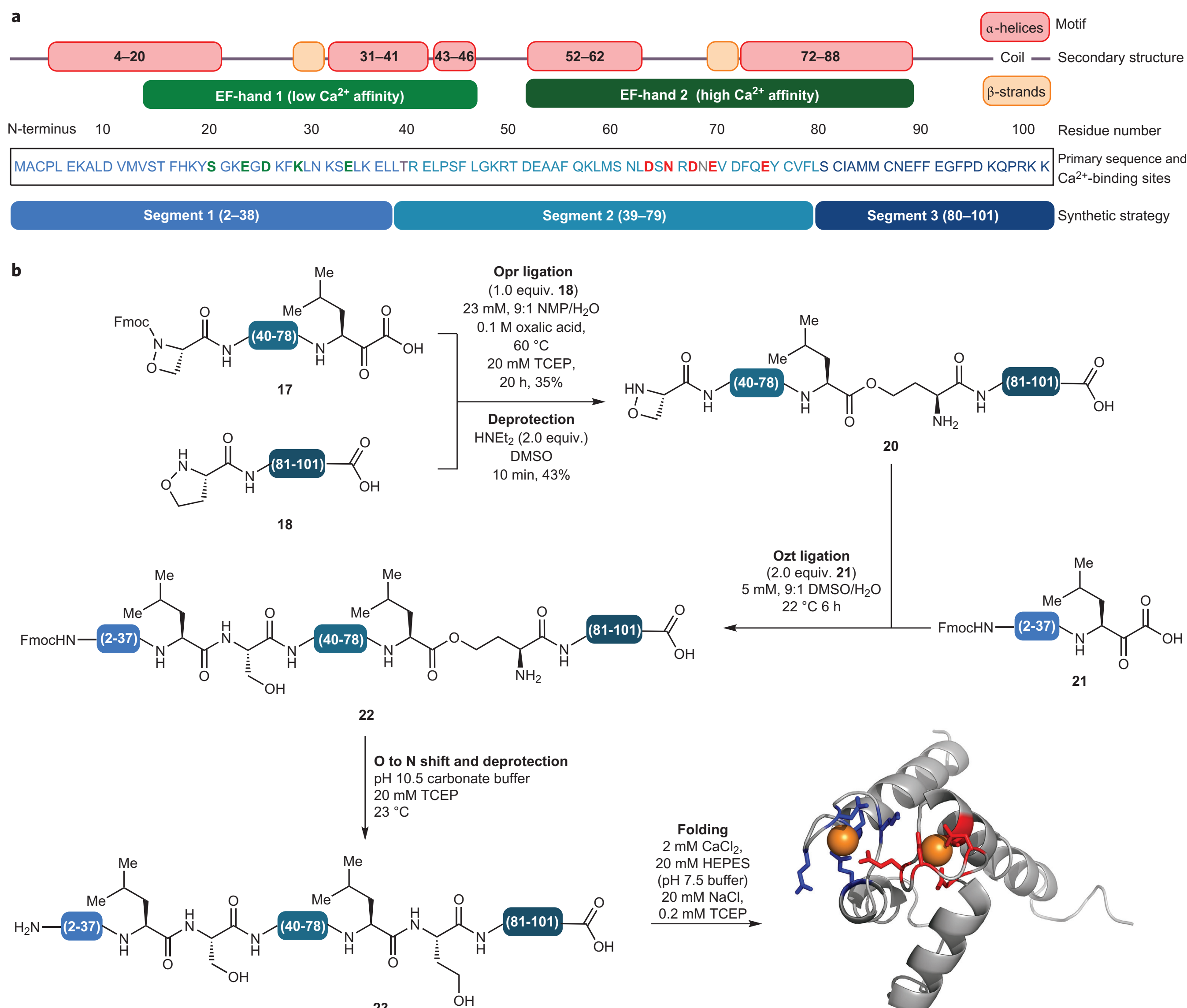


Figure 3 | Structure and synthesis of S100A4. a, Structural motifs, metal-binding sites and synthetic strategy. **b**, Assembly of S100A4 by the combination of KAHA ligation with 5-Opr and Ozt. The final ligation was carried out at 5 mM and 22 °C. TCEP, tris(2-carboxyethyl)phosphine; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid.

and O-to-N acyl shift to give 23, which was folded in the presence of Ca^{2+}. The folding and metal binding was confirmed by circular dichroism (CD) spectroscopy (Fig. 4). Evaluation of this synthetic S100A4 in an adipocyte proliferation assay confirmed its biological activity.

In conclusion, the synthesis of an enantioenriched oxazetidine amino acid—formally an oxidized form of serine—enables the first native amide-forming ligation reaction that compares to the venerable NCL in both reaction rate[26] and chemoselectivity. By operating at prevalent serine (6.8% abundance)[27,28], rather than cysteine (1.8% abundance), this ligation offers greater flexibility in synthetic planning. Serine residues can also easily replace many other residues without affecting protein folding or function. We have recently introduced solid supported linkers for Fmoc-SPPS that afford peptide α-ketoacids directly upon resin cleavage[8]. As these resins and the Fmoc-(*S*)-Ozt monomer become widely available, we expect that this variant of the KAHA ligation will emerge as a leading approach to chemical protein synthesis due to its ease of execution, traceless ligation of the peptide segments and facile implementation of multiple segment couplings.

Received 1 March 2015; accepted 12 May 2015; published online 22 June 2015

References

1. Kent, S. B. H. Bringing the science of proteins into the realm of organic chemistry: total chemical synthesis of SEP (synthetic erythropoiesis protein). *Angew. Chem. Int. Ed.* **52,** 11988–11996 (2013).
2. Nilsson, B. L., Soellner, M. B. & Raines, R. T. Chemical synthesis of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **34,** 91–118 (2005).
3. Hackenberger, C. P. R., Bode, J. W. & Schwarzer, D. in *Amino Acids, Peptides and Proteins in Organic Chemistry: Building Blocks, Catalysis and Coupling Chemistry* Vol. 3 (ed Hughes, A. B.) 445–493 (Wiley-VCH, 2010).
4. Dawson, P., Muir, T., Clark-Lewis, I. & Kent, S. Synthesis of proteins by native chemical ligation. *Science* **266,** 776–779 (1994).
5. Pattabiraman, V. R. & Bode, J. W. Rethinking amide bond synthesis. *Nature* **480,** 471–479 (2011).
6. Bode, J. W., Fox, R. M. & Baucom, K. D. Chemoselective amide ligations by decarboxylative condensations of *N*-alkylhydroxylamines and α-ketoacids. *Angew. Chem. Int. Ed.* **45,** 1248–1252 (2006).
7. Pattabiraman, V. R., Ogunkoya, A. O. & Bode, J. W. Chemical protein synthesis by chemoselective α-ketoacid-hydroxylamine (KAHA) ligations with 5-oxaproline. *Angew. Chem. Int. Ed.* **51,** 5114–5118 (2012).

© 2015 Macmillan Publishers Limited. All rights reserved

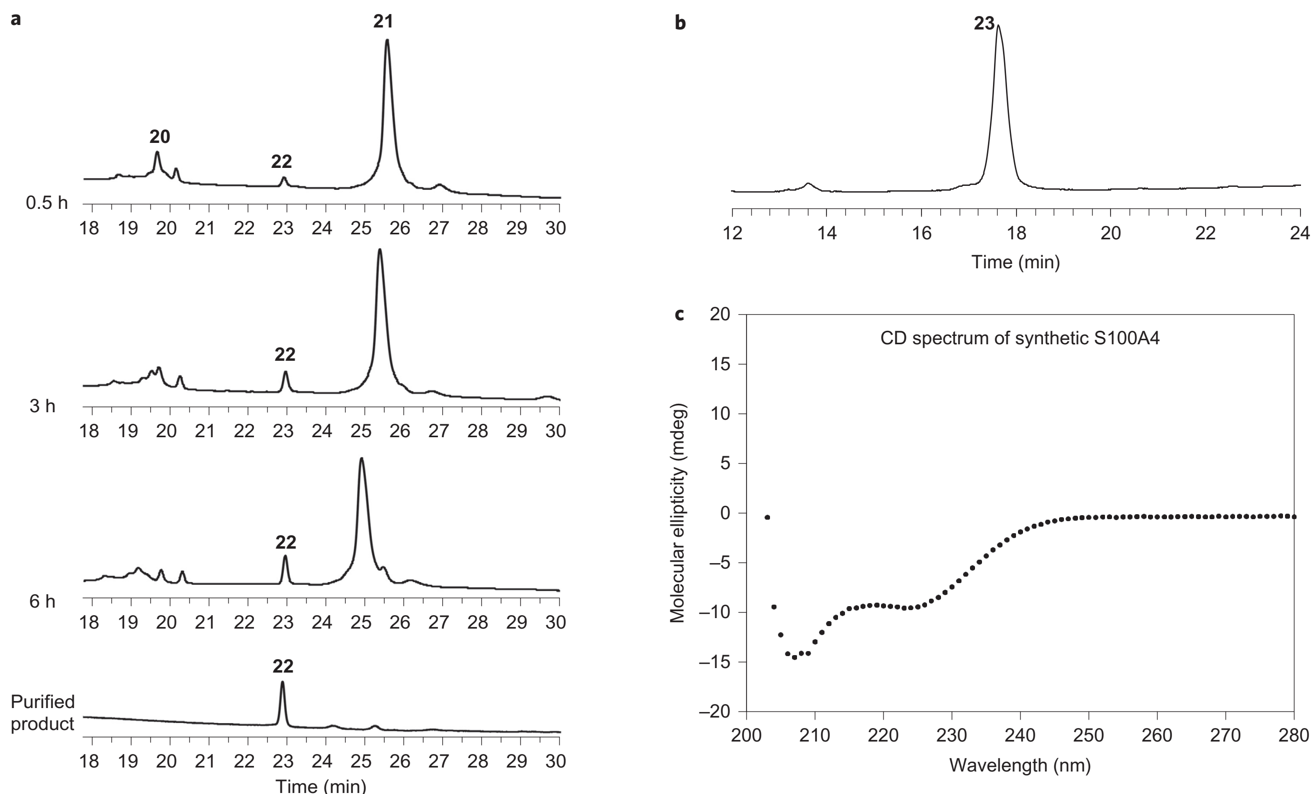


Figure 4 | Monitoring of the coupling and folding processes in the synthesis of S100A4. a, Oxezetadine ligation of 20 and 21. To facilitate monitoring, the Fmoc group was retained on the N-terminus of 21, which displays greatly increased absorbance compared to 20. **b**, HPLC of final coupled and purified product 23. **c**, CD spectrum of folded synthetic protein S100A4.

8. Wucherpfennig, T. G., Pattabiraman, V. R., Limberg, F. R. P., Ruiz-Rodríguez, J. & Bode, J. W. Traceless preparation of C-terminal α-ketoacids for chemical protein synthesis by α-ketoacid–hydroxylamine ligation: synthesis of SUMO2/3. *Angew. Chem. Int. Ed.* **53,** 12248–12252 (2014).
9. Wucherpfennig, T. G., Rohrbacher, F., Pattabiraman, V. R. & Bode, J. W. Formation and rearrangement of homoserine depsipeptides and depsiproteins in the α-ketoacid–hydroxylamine ligation with 5-oxaproline. *Angew. Chem Int. Ed.* **53,** 12244–12247 (2014).
10. Magers, D. H. & Davis, S. R. Ring strain in the oxazetidines. *J. Mol. Struc. Theochem*. **487,** 205–210 (1999).
11. Florio, S., Capriati, V., & Luisi, R. in *Comprehensive Heterocyclic Chemistry III* Vol. 2 (eds Katritzky, A. R., Ramsden, C. A., Scriven, E. F. V. & Taylor, R. J. K.) Ch. 14, 689–711 (Elsevier, 2008).
12. Snider, B. B. & Duvall, J. R. Synthesis of the 4-methyl-1,2-oxazetidine-4-carboxylic acid moiety of the originally proposed halipeptin A and B structures. *Tetrahedron Lett.* **44,** 3067–3070 (2003).
13. Denicola, A., Einhorn, C., Einhorn, J. & Luche, J. L. Intramolecular alkylation of carboxylic-acids—application to the synthesis of Boc-protected cyclic amino-acids. *J. Chem. Soc. Chem. Commun*. 879–880 (1994).
14. Casadei, M. A., Galli, C. & Mandolini, L. Ring-closure reactions. 22. Kinetics of cyclization of diethyl (ω-bromoalkyl)malonates in the range of 4- to 21-membered rings. Role of ring strain. *J. Am. Chem. Soc.* **106,** 1051–1056 (1984).
15. Baldwin, J. E. Rules for ring closure. *Chem. Commun*. 734–736 (1976).
16. Alabugin, I. V. & Gilmore, K. Finding the right path: Baldwin 'rules for ring closure' and stereoelectronic control of cyclizations. *Chem. Commun.* **49,** 1124–11250 (2013).
17. Tanino, K. *et al.* Total synthesis of Solanoeclepin A. *Nature Chem*. **3,** 484–488 (2011).
18. Medjahdi, M., Gonzalez-Gomez, J. C., Foubelo, F. & Yus, M. Stereoselective synthesis of azetidines and pyrrolidines from *N-tert*-butylsulfonyl(2-aminoalkyl)oxiranes. *J. Org. Chem*. **74,** 7859–7865 (2009).
19. Davies, S. G., Jones, S., Sanz, M. A., Teixeira, F. C. & Fox, J. F. A novel [2,3] intramolecular rearrangement of *N*-benzyl-*O*-allylhydroxylamines. *Chem. Commun*. 2235–2236 (1998).
20. Wuts, P. G. M. & Greene, T. W. in *Greene's Protective Groups in Organic Synthesis* (Wiley, 2006).
21. Vallely, K. M. *et al.* Solution structure of human Mts1 (S100A4) as determined by NMR spectroscopy. *Biochemistry* **41,** 12670–12680 (2002).
22. Kiss, B. *et al.* Crystal structure of the S100A4-nonmuscle myosin IIA tail fragment complex reveals an asymmetric target binding mechanism. *Proc. Natl Acad. Sci. USA* **109,** 6048–6053 (2012).
23. Schäfer, B. W. & Heizmann, C. W. The S100 family of EF-hand calcium-binding proteins: functions and pathology. *Trends Biochem. Sci*. **21,** 134–140 (1996).
24. Sohma, Y., Sasaki, M., Hayashi, Y., Kimura, T. & Kiso, Y. Novel and efficient synthesis of difficult sequence-containing peptides through O–N intramolecular acyl migration reaction of O-acyl E isopeptides. *Chem. Commun.* 124–125 (2004).
25. Avital-Shmilovici, M. *et al.* Fully convergent chemical synthesis of ester insulin: determination of the high resolution X-ray structure by racemic protein crystallography. *J. Am. Chem. Soc.* **135,** 3173–3185 (2013).
26. Kalia, J. & Raines, R. T. Advances in bioconjugation. *Curr. Org. Chem*. **14,** 138–147 (2010).
27. Li, X., Lam, H. Y., Zhang, Y. & Chan, C. K. Salicylaldehyde ester-induced chemoselective peptide ligations: enabling generation of natural peptidic linkages at the serine/threonine sites. *Org. Lett.* **12,** 1724–1727 (2010).
28. Zhang, Y., Xu, C., Lam, H. Y., Lee, C. L., & Li, X. Protein chemical synthesis by serine and threonine ligation. *Proc. Natl Acad. Sci. USA* **110,** 6657–6662 (2012).

Acknowledgements

This work was supported by the Swiss National Science Foundation (200020_150073) and ETH Zürich. F. Thuaud, S. Baldauf and M. Dao are thanked for contributions to the synthesis of compound 1, V. Pattabiraman for discussions, F. Saito for advice on kinetics and C. Wolfrum for biological evaluation.

Author contributions

J.W.B. and I.P. contributed equally to the design of the study. J.B., with contributions from I.P., wrote the paper. I.P performed the experiments and wrote the Supplementary Information.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.W.B.

Competing financial interests

The authors declare no competing financial interests.

© 2015 Macmillan Publishers Limited. All rights reserved