LETTER

doi:10.1038/nature09981

Enzyme-catalysed [4+2] cycloaddition is a key step in the biosynthesis of spinosyn A

Hak Joong Kim[1], Mark W. Ruszczycky[2]*, Sei-hyun Choi[1]*, Yung-nan Liu[2] & Hung-wen Liu[1,2]

The Diels–Alder reaction is a [4+2] cycloaddition reaction in which a cyclohexene ring is formed between a 1,3-diene and an electron-deficient alkene via a single pericyclic transition state[1]. This reaction has been proposed as a key transformation in the biosynthesis of many cyclohexene-containing secondary metabolites[2–5]. However, only four purified enzymes have thus far been implicated in bio-transformations that are consistent with a Diels–Alder reaction, namely solanapyrone synthase[6], LovB[7,8], macrophomate synthase[9,10], and riboflavin synthase[11,12]. Although the stereochemical outcomes of these reactions indicate that the product formation could be enzyme-guided in each case, these enzymes typically demonstrate more than one catalytic activity, leaving their specific influence on the cycloaddition step uncertain. In our studies of the biosynthesis of spinosyn A, a tetracyclic polyketide-derived insecticide from *Saccharopolyspora spinosa*[13,14], we identified a cyclase, SpnF, that catalyses a transannular [4+2] cycloaddition to form the cyclohexene ring in spinosyn A. Kinetic analysis demonstrates that SpnF specifically accelerates the ring formation reaction with an estimated 500-fold rate enhancement. A second enzyme, SpnL, was also identified as responsible for the final cross-bridging step that completes the tetracyclic core of spinosyn A in a manner consistent with a Rauhut–Currier reaction[15]. This work is significant because SpnF represents the first example characterized *in vitro* of a stand-alone enzyme solely committed to the catalysis of a [4+2] cycloaddition reaction. In addition, the mode of formation of the complex perhydro-*as*-indacene moiety in spinosyn A is now fully established.

Spinosyn A (**1**), an active ingredient of several highly effective and environmentally benign commercial insecticides, has a complex aglycone structure comprising a perhydro-*as*-indacene moiety fused to a 12-membered macrolactone[13,14]. How this tetracyclic ring system is biosynthesized has been a subject of much speculation[16–18]. Attention has largely focused on the construction of the cyclohexene ring due to the potential involvement of an enzyme that catalyses the [4+2] cycloaddition, which if concerted would represent a so-called 'Diels–Alderase'. Four genes in the spinosyn A biosynthetic gene cluster of *S. spinosa*—*spnF*, *spnJ*, *spnL* and *spnM*—were proposed to convert product (**2**) of the polyketide synthase (PKS) to the tetracyclic aglycone (**4**) (see Fig. 1)[16]. The gene product of *spnJ*, which is a flavin-dependent dehydrogenase, was recently demonstrated to catalyse oxidation of the 15-OH of **2** to form the keto-intermediate **3** (ref. 19). However, the functions of the enzymes encoded by the remaining three genes, *spnF*, *spnL* and *spnM*, which show significant sequence similarity to lipases (SpnM) and *S*-adenosyl-L-methionine (SAM)-dependent methyltransferases (SpnF and SpnL)[16], remain elusive.

To investigate the functions of SpnF, SpnL and SpnM, the corresponding genes were heterologously overexpressed in *Escherichia coli* BL21(DE3)* and their products purified as *N*-terminal His_6-tagged proteins (>95% purity). Each of the purified enzymes was tested for activity with **3**. As shown in Fig. 2A, neither SpnF nor SpnL processes **3**. In contrast, complete conversion of **3** to a new product was observed after a 2-h incubation with SpnM. NMR and mass spectrometry analysis of this new product (**8**) revealed a 15,6,5-tricyclic skeleton. Further investigation of the reaction time-course led to the discovery of a transient intermediate (Fig. 2D, orange peak), which was identified by spectral analysis as the monocyclic macrolactone **5**.

These findings indicated that the SpnM-catalysed conversion of **3** to **8** might be a two-step process involving 1,4-dehydration of **3** followed by a transannular [4+2] cycloaddition between the $\Delta^{11,12}$-alkene and the conjugated $\Delta^{4,5}$,$\Delta^{6,7}$-diene of intermediate **5** to form the cyclohexene moiety in **8** (Fig. 4). To investigate whether SpnM is indeed a

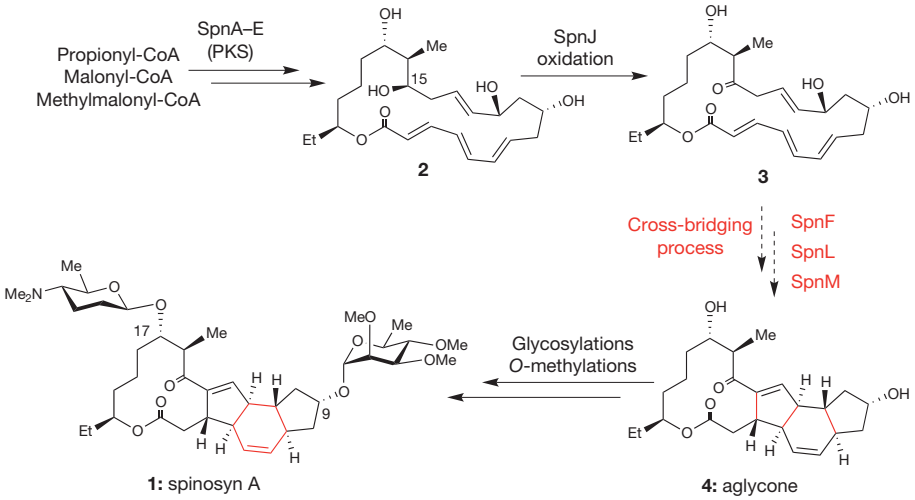

Figure 1 | Proposed spinosyn biosynthetic pathway.

[1]Department of Chemistry and Biochemistry, University of Texas at Austin, Austin, Texas 78712, USA. [2]Division of Medicinal Chemistry, College of Pharmacy, University of Texas at Austin, Austin, Texas 78712, USA.
*These authors contributed equally to this work.

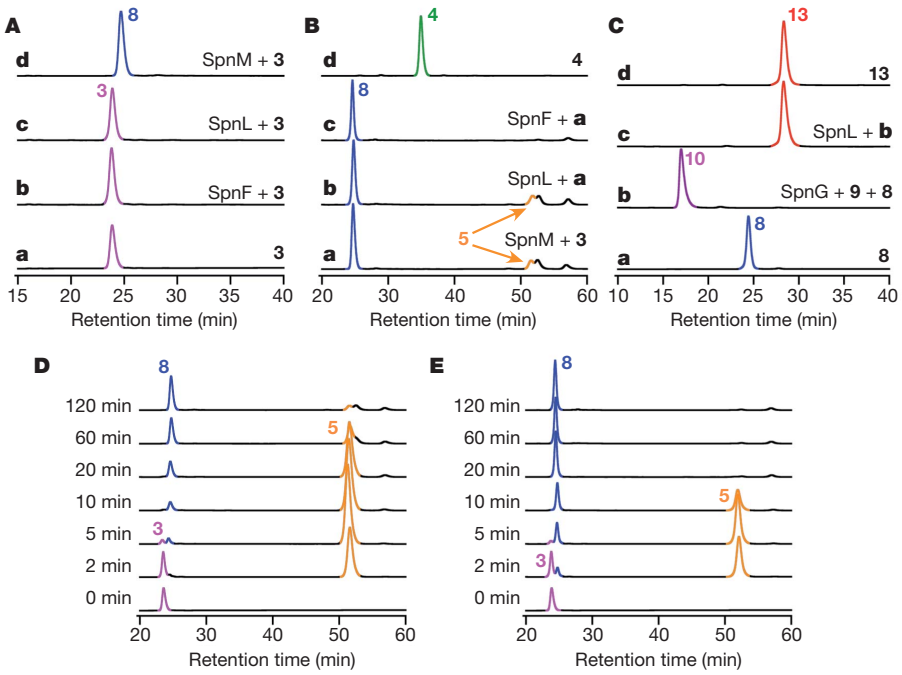

Figure 2 | HPLC analysis showing the reactions catalysed by SpnF, SpnL and SpnM. A, Incubation of 2 mM **3** for 2 h alone (**a**), and with either 1.25 μM SpnF (**b**), 1.25 μM SpnL (**c**), or 1.25 μM SpnM (**d**). **B**, Incubation of 2 mM **8** (generated *in situ* from **3**) for 2 h either alone (**a**), with 15 μM SpnL (**b**), or with 15 μM SpnF (**c**). Trace (**d**) is the authentic standard of **4**. **C**, Incubation of 1 mM **8** for 2 h either alone (**a**); with both 20 μM SpnG and 1.5 mM **9** (**b**); or with 15 μM SpnL, 20 μM SpnG and 1.5 mM **9** (**c**). Trace (**d**) is the authentic standard of **13**. **D**, Time-course of 2 mM **3** in the presence of 1.25 μM SpnM. **E**, Time course of 2 mM **3** in the presence of both 1.25 μM SpnM and 10 μM SpnF.

cyclase of dual function catalysing both the dehydration and cyclization steps, the dependence of the rate of each step on the concentration of SpnM was determined. In this experiment the formation and consumption of **5**, which respectively reflect the dehydration and cyclization steps, was monitored by high-performance liquid chromatography (HPLC) as a function of time.

As shown in Fig. 3a, rate enhancement of the dehydration step was observed with the increase in SpnM concentration, whereas the rate of cyclization was unaffected. The full time courses of the formation and decay of **5** were analysed using numerically integrated coupled rate equations based on different kinetic models (see Supplementary Information Section 5)[20]. The observed data was fit best to the integrated rate equations (1a) and (1b), which model Michaelis–Menten kinetics for the dehydration step (V_{SpnM} and K_{SpnM}) and first order kinetics (k_{non}) for the cyclization step. Correlation of the fitted parameters versus the concentration of SpnM (Fig. 3b) reveals a significant dependence of V_{SpnM}, but not K_{SpnM} or k_{non}, on SpnM concentration. This analysis indicates that SpnM functions only as a dehydratase, and its product (**5**) can cyclize nonenzymatically to **8** with a first order rate constant of approximately 0.03 min^{-1}.

Having established that SpnM functions as a dehydratase, whereas the [4+2] cycloaddition can proceed nonenzymatically, we next considered the possible involvement of the two remaining gene products,

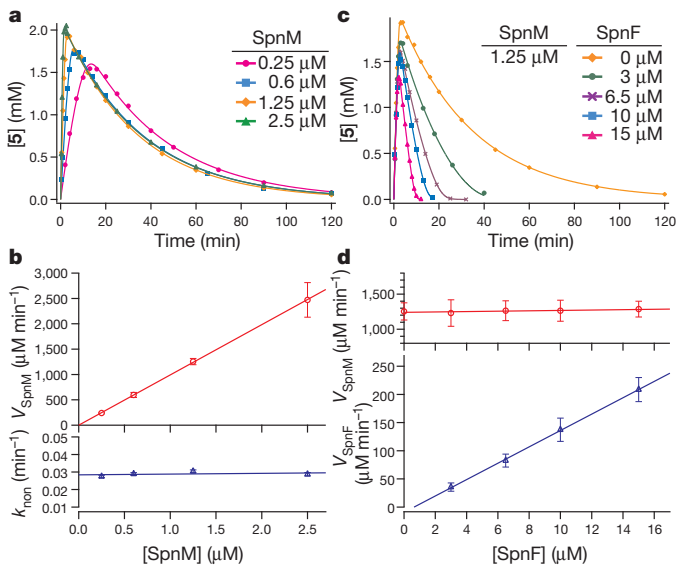

Figure 3 | Kinetic analysis demonstrating that SpnM and SpnF, respectively, catalyse the dehydration and cyclization steps of macrolactone 3. a, **c**, Formation and consumption of **5** was monitored by HPLC and the concentration ([**5**]) plotted versus time. Each reaction mixture initially contained 2 mM **3**, and the indicated amounts of SpnM and SpnF. **a**, Variable SpnM with no SpnF and fits based on the rate equations (1).

$$\frac{d[\mathbf{3}]}{dt} = -\frac{V_{SpnM}[\mathbf{3}]}{K_{SpnM}+[\mathbf{3}]} \quad (1a)$$

$$\frac{d[\mathbf{5}]}{dt} = \frac{V_{SpnM}[\mathbf{3}]}{K_{SpnM}+[\mathbf{3}]} - k_{non}[\mathbf{5}] \quad (1b)$$

b, Fitted parameters V_{SpnM} and k_{non} versus [SpnM]. The measured turnover number for SpnM, $k_{cat,SpnM}$, is 1,020 ± 57 min^{-1} whereas the first order rate constant for nonenzymatic cyclization of **5**, k_{non}, is 0.0288 ± 0.00041 min^{-1}. The apparent Michaelis constant for SpnM, K_{SpnM}, is 380 ± 51 μM. **c**, Variable SpnF at a fixed concentration of SpnM and fits based on rate equations (2).

$$\frac{d[\mathbf{3}]}{dt} = -\frac{V_{SpnM}[\mathbf{3}]}{K_{SpnM}+[\mathbf{3}]} \quad (2a)$$

$$\frac{d[\mathbf{5}]}{dt} = \frac{V_{SpnM}[\mathbf{3}]}{K_{SpnM}+[\mathbf{3}]} - k_{non}[\mathbf{5}] - \frac{V_{SpnF}[\mathbf{5}]}{K_{SpnF}+[\mathbf{5}]} \quad (2b)$$

d, Correlation of V_{SpnM} and V_{SpnF} versus [SpnF]. The turnover number for SpnF, $k_{cat,SpnF}$, is 14 ± 1.6 min^{-1}. The apparent Michaelis constant for SpnF, K_{SpnF}, which is independent of SpnF concentration, is 120 ± 46 μM. All values and error bars are ± s.e.

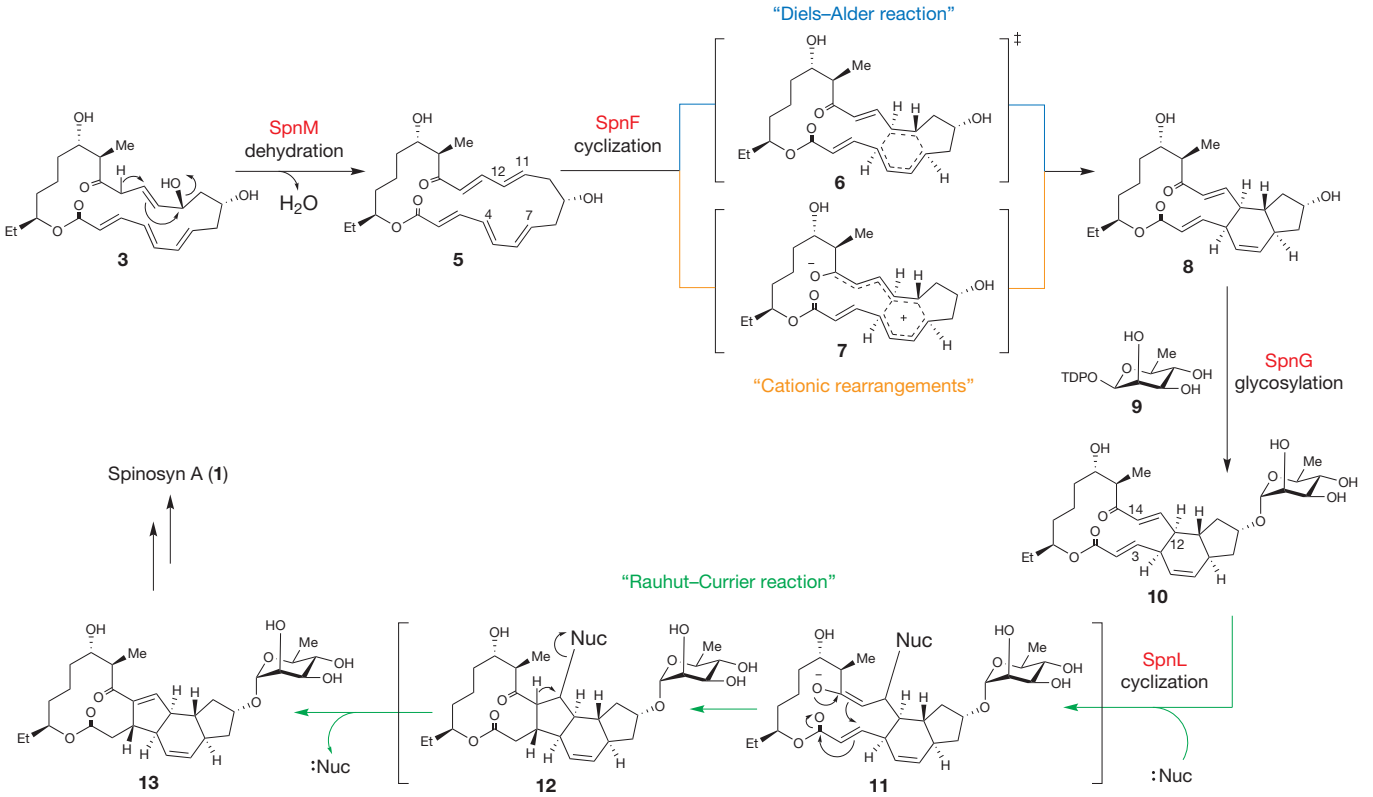

Figure 4 | The spinosyn aglycone biosynthetic pathway. SpnM catalyses a dehydration reaction to convert **3** to **5**, and SpnF subsequently catalyses cyclization of **5** to afford **8**. The resulting tricyclic macrolactone **8** is then modified with a rhamnose moiety at the C-9 hydroxyl group by SpnG rather than being converted directly to the aglycone **4** (see Fig. 1) as previously thought. SpnL completes the cross-bridging process by interlinking the C-3 and C-14 carbon centres of **10** to produce the tetracyclic nucleus (**13**) of spinosyn A.

SpnF and SpnL, in the final C-C bond formation between C-3 and C-14. Surprisingly, incubation of SpnF or SpnL with **8**, which was generated *in situ* from **3** in the presence of SpnM, did not produce the aglycone **4** (Fig. 2B). However, when SpnF was added to the assay mixture, a change in the product distribution was noted including rapid disappearance of **5** (Fig. 2B, trace **c**). As shown in Fig. 2E, the consumption of **5** with concomitant formation of a product having a retention time consistent with **8** was completed in 20 min in the presence of 10 μM SpnF, instead of requiring more than 2 h in its absence (Fig. 2D). The structure of this product was established to be **8** by NMR analysis. Additional controls, including examining the effects of denatured SpnF as well as mutated SpnF on the rate of cyclization, were also performed. In all cases only native (His_6-tagged) SpnF led to appreciably increased rates of cyclization (see Supplementary Information Section 4.4). These results clearly indicate that SpnF is responsible for the observed rate enhancement of the cyclization of **5** to **8**.

To quantify the effect of SpnF on the rate of the cyclization step, the production and consumption of **5** was again monitored by HPLC, this time as a function of SpnF concentration while keeping that of SpnM fixed. As shown in Fig. 3c, the rate of conversion of **5** to **8** is clearly enhanced as the concentration of SpnF increases. The data was fit best using the coupled rate equations (2a) and (2b), where the cyclization event is modelled as the sum of a first order and a Michaelis–Menten process. In these fits all parameters including the initial concentration of **3** were allowed to float except for k_{non}, which was fixed at the fitted value obtained in the absence of SpnF. Correlation of the fitted parameters with SpnF concentration indicates a significant dependence of V_{SpnF} on SpnF concentration (Fig. 3d), which is not true for the V_{SpnM} parameter. These results firmly establish that SpnF is a cyclase catalysing the conversion of **5** to **8** with an apparent k_{cat} of $14 \pm 1.6\ min^{-1}$ ($\pm$ s.e.) for an estimated rate enhancement ($k_{cat,SpnF}$ versus k_{non}) of approximately 500-fold.

This result left *spnL* as the only gene without an assigned function, and it was proposed that SpnL has a role in the transannular cyclization reaction between C-3 and C-14. However, the observed inability of SpnL to convert **8** to the anticipated product (**4**) (Fig. 2B, trace **b**) prompted us to reconsider the sequence of events of the proposed biosynthetic pathway in Fig. 1. Previously, glycosylations had been thought to occur on the tetracyclic aglycone (**4**), because transfer of the rhamnose moiety from thymidine diphosphate-β-L-rhamnose (TDP-β-L-rhamnose, **9**) to 9-OH of **4** by the rhamnosyltransferase, SpnG, had been demonstrated[21,22]. However, this observation did not necessarily rule out an alternative pathway in which rhamnosylation of **8** precedes C-3/C-14 cross-bridging, because glycosyltransferases involved in the biosynthesis of secondary metabolites are known to be promiscuous with regard to their substrate specificity[23,24].

To test this possibility, the ability of SpnG to accept **8** as a substrate was investigated. As shown in Fig. 2C (trace b), incubation of **8** and **9** with SpnG resulted in the disappearance of **8** with the concomitant appearance of a new peak (retention time of 17 min). This new product was identified as **10** by NMR and mass spectrometry analysis (Fig. 4). Upon addition of SpnL to this reaction mixture, the peak at 17 min disappeared and a new peak appeared at 27 min (Fig. 2C, trace c). Both transformations were highly efficient. Identification of the new product as **13** was confirmed by high-resolution mass spectrometry analysis and HPLC co-elution with an authentic sample of **13**. These results strongly suggest that **10** rather than **8** is the biological substrate for SpnL, which catalyses the final cyclization step to generate the perhydro-*as*-indacene core.

The mechanisms by which SpnF and SpnL catalyse their respective cyclization reactions are a point of interest. The SpnF-catalysed *endo*-mode *syn*-addition of an alkenyl to a dienyl functionality seems consistent with a Diels–Alder reaction; however, confirmation of this hypothesis will require demonstrating that the reaction progresses

through a single pericyclic transition state such as **6**. Therefore, a stepwise [4+2] cycloaddition mechanism, for example, one involving a dipolar intermediate such as **7**, cannot at present be ruled out. In contrast, the C-C bond formation catalysed by SpnL may involve a Rauhut–Currier mechanism[15] consistent with the observation that **10** is susceptible towards nucleophilic addition by a thiol (see Supplementary Information Section 3.10) forming a covalent adduct that may be structurally analogous to **11** or **12** (see Fig. 4), although the specific site of attack remains unknown. Whereas these mechanistic proposals are at present speculative, it is worth noting that Roush and co-workers were able to accomplish their non-enzymatic total synthesis of spinosyn A by exploiting both the transannular Diels–Alder and Rauhut–Currier reactions in an analogous fashion[25]. This precedent in chemical synthesis certainly substantiates the feasibility of the mechanisms proposed for the SpnF- and SpnL-catalysed reactions.

In summary, the biosynthetic pathway for spinosyn A is now fully established (Fig. 4), with the specific functions of SpnM as a dehydratase and SpnF as well as SpnL as the two cyclases in the cross-bridging steps biochemically determined. SpnF represents the first enzyme for which specific acceleration of a [4+2] cycloaddition reaction has been experimentally verified as its only known function. It stands in contrast to macrophomate synthase, for which evidence has been provided suggesting a tandem Michael–aldol reaction mechanism[26,27], as well as the multifunctional solanapyrone synthase, LovB and riboflavin synthase, which participate in hydroxyl oxidation[5,6], polyketide synthesis[7,8], and hydride transfer[12], respectively, in addition to the [4+2] cycloaddition reactions, the concertedness of which have yet to be verified. For this reason, the SpnF reaction provides a unique system for detailed mechanistic investigation of enzyme-catalysed [4+2] cycloadditions and the existence of a bona fide Diels–Alderase.

METHODS SUMMARY

All proteins used in this work were overexpressed in *E. coli* BL21(DE3)* (Invitrogen) and purified by Ni-NTA (Qiagen) affinity chromatography. Specifically, SpnF was co-overexpressed with the chaperone protein pair, GroEL/ES, to improve its solubility. Because overexpression of the protein encoded by the originally assigned *spnM* gene[16] failed to afford an active soluble protein, the gene sequence was re-examined and revised to include 204 additional base pairs (Supplementary Fig. 2). Overexpression of the revised *spnM* gene produced an active enzyme with significantly improved protein yield. All enzyme reaction products (**5**, **8**, **10** and **13**) were extracted with ethyl acetate or chloroform and purified using silica gel column chromatography or HPLC. Their structures were characterized by 1D- and 2D-NMR spectroscopy and/or high-resolution mass spectrometry. In particular, the stereochemistry of **8** was assigned based on its ^{1}H-^{1}H nuclear Overhauser enhancement spectroscopy (NOESY) spectrum. All substrate specificity and time-course assays were run in 50 mM Tris-HCl buffer (pH 8) at 30 °C. Reaction aliquots were quenched with an excess volume of ethanol after a given incubation time and centrifuged to remove protein. The supernatant was then analysed by reverse phase HPLC with detection by ultraviolet absorbance at 254 nm (Fig. 2) or 280 nm (Fig. 3). Time course assays also included *p*-methoxyacetophenone as an internal standard to normalize the peak areas corresponding to **5**. Numerical integration of equations (1) and (2) used the fourth order Runge–Kutta algorithm[28] following non-dimensionalization of substrate concentration. The resulting simulated progress curves were fit using steepest descent[29] directly to full time-courses of normalized substrate concentration obtained via the HPLC discontinuous assay to provide the kinetic parameters and a concentration normalization factor. Further detail regarding experimental procedures as well as data fitting and analysis is described in the Supplementary Information.

Received 29 March 2010; accepted 3 March 2011.

1. Huisgen, R. Cycloadditions — definition, classification, and characterization. *Angew. Chem. Int. Edn Engl.* **7,** 321–328 (1968).
2. Stocking, E. M. & Williams, R. M. Chemistry and biology of biosynthetic Diels–Alder reactions. *Angew. Chem. Int. Ed.* **42,** 3078–3115 (2003).
3. Oikawa, H. Involvement of the Diels–Alderases in the biosynthesis of natural products. *Bull. Chem. Soc. Jpn.* **78,** 537–554 (2005).
4. Kelly, W. L. Intramolecular cyclizations of polyketide biosynthesis: mining for a ''Diels–Alderase?''. *Org. Biomol. Chem.* **6,** 4483–4493 (2008).
5. Oikawa, H. in *Comprehensive Natural Products II, Chemistry and Biology* (eds Mander, L. & Liu, H.-w.) Vol. 8 277–314 (Elsevier, 2010).
6. Oikawa, H., Katayama, K., Suzuki, Y. & Ichihara, A. Enzymatic activity catalyzing *exo*-selective Diels–Alder reaction in solanapyrone biosynthesis. *J. Chem. Soc. Chem. Comm.* 1321–1322 (1995).
7. Auclair, K. *et al.* Lovastatin nonaketide synthase catalyzes an intramolecular Diels–Alder reaction of a substrate analogue. *J. Am. Chem. Soc.* **122,** 11519–11520 (2000).
8. Ma, S. M. *et al.* Complete reconstitution of a highly reducing iterative polyketide synthase. *Science* **326,** 589–592 (2009).
9. Watanabe, K., Mie, T., Ichihara, A., Oikawa, H. & Honma, M. Detailed reaction mechanism of macrophomate synthase. *J. Biol. Chem.* **275,** 38393–38401 (2000).
10. Ose, T. *et al.* Insight into a natural Diels–Alder reaction from the structure of macrophomate synthase. *Nature* **422,** 185–189 (2003).
11. Eberhardt, S. *et al.* Domain structure of riboflavin synthase. *Eur. J. Biochem.* **268,** 4315–4323 (2001).
12. Kim, R.-R. *et al.* Mechanistic insights on riboflavin synthase inspired by selective binding of the 6,7-dimethyl-8-ribityllumazine exomethylene anion. *J. Am. Chem. Soc.* **132,** 2983–2990 (2010).
13. Kirst, H. A. *et al.* A83543A-D, Unique fermentation-derived tetracyclic macrolides. *Tetrahedr. Lett.* **32,** 4839–4842 (1991).
14. Kirst, H. A. The spinosyn family of insecticides: realizing the potential of natural products research. *J. Antibiot.* **63,** 101–111 (2010).
15. Aroyan, C. E., Dermenci, A. & Miller, S. J. The Rauhut–Currier reaction: a history and its synthetic application. *Tetrahedron* **65,** 4069–4084 (2009).
16. Waldron, C. *et al.* Cloning and analysis of the spinosad biosynthetic gene cluster of *Saccharopolyspora spinosa*. *Chem. Biol.* **8,** 487–499 (2001).
17. Martin, C. J. *et al.* Heterologous expression in *Saccharopolyspora erythraea* of a pentaketide synthase derived from the spinosyn polyketide synthase. *Org. Biomol. Chem.* **1,** 4144–4147 (2003).
18. Oikawa, H. Biosynthesis of structurally unique fungal metabolite GKK1032A_2: indication of novel carbocyclic formation mechanism in polyketide biosynthesis. *J. Org. Chem.* **68,** 3552–3557 (2003).
19. Kim, H. J., Pongdee, R., Wu, Q., Hong, L. & Liu, H.-w. The biosynthesis of spinosyn in *Saccharopolyspora spinosa*: synthesis of the cross-bridging precursor and identification of the function of SpnJ. *J. Am. Chem. Soc.* **129,** 14582–14584 (2007).
20. Frieden, C. Analysis of kinetic data: practical applications of computer simulation and fitting programs. *Methods Enzymol.* **240,** 311–322 (1994).
21. Chen, Y., Lin, Y., Tsai, K. & Chiu, H. Functional characterization and substrate specificity of spinosyn rhamnosyltransferase by *in vitro* reconstitution of spinosyn biosynthetic enzymes. *J. Biol. Chem.* **284,** 7352–7363 (2009).
22. Kim, H. J. *et al.* Biosynthesis of spinosyn in *Saccharopolyspora spinosa*: synthesis of permethylated rhamnose and characterization of the functions of SpnH, SpnI, and SpnK. *J. Am. Chem. Soc.* **132,** 2901–2903 (2010).
23. Thibodeaux, C. J., Melançon, C. E. & Liu, H.-w. Unusual sugar biosynthesis and natural product glycodiversification. *Nature* **446,** 1008–1016 (2007).
24. Thibodeaux, C. J., Melançon, C. E. & Liu, H.-w. Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew. Chem. Int. Ed.* **47,** 9814–9859 (2008).
25. Mergott, D. J., Frank, S. A. & Roush, W. R. Total synthesis of (−)-spinosyn A. *Proc. Natl Acad. Sci. USA* **101,** 11955–11959 (2004).
26. Guimarães, C. R. W., Udier-Blagović, M. & Jørgensen, W. L. Macrophomate synthase: QM/MM simulations address the Diels–Alder versus Michael–aldol reaction mechanism. *J. Am. Chem. Soc.* **127,** 3577–3588 (2005).
27. Serafimov, J. M., Gillingham, D., Kuster, S. & Hilvert, D. The putative Diels–Alderase macrophomate synthase is an efficient aldolase. *J. Am. Chem. Soc.* **130,** 7798–7799 (2008).
28. Tenenbaum, M. & Pollard, H. *Ordinary Differential Equations* (Dover, 1985).
29. Draper, N. R. & Smith, H. *Applied Regression Analysis* 3rd edn (Wiley-Interscience, 1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. Whitman for review of the manuscript, L. Hong for carrying out the early cloning work, B. Shoulders and S. Sorey for assistance with the interpretation of the NMR spectra, and S. Mansoorabadi and E. Isiorho for discussions on the reaction mechanisms and structural modelling of SpnF. This work is supported in part by grants from the National Institutes of Health (GM035906, F32AI082906), the Texas Higher Education Coordination Board (ARP-003658-0093-2007), and the Welch Foundation (F-1511).

Author Contributions H.-w.L. provided the scientific direction and the overall experimental design for the studies. H.J.K. designed and performed most of the experiments. S.-h.C. participated in the chemical synthesis of the substrate for SpnJ (**2**) and the characterization of the structures of the enzyme reaction products. M.W.R. analysed the kinetic experiment data. Y.-n.L. carried out the mutation studies of SpnF. H.J.K., M.W.R. and H.-w.L. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to H.-w.L. (h.w.liu@mail.utexas.edu).