CHEMISTRY
A EUROPEAN JOURNAL

FULL PAPER

Preparation and characterization of a bifunctional aldolase/kinase enzyme. A more efficient biocatalyst for C-C bond formation

Laura Iturrate,[a] Israel Sánchez-Moreno,[a] Isabel Oroz-Guinea,[a] Jesús Pérez-Gil[b] and Eduardo García-Junceda*[a]

Dedication ((optional))

Abstract: A bifunctional aldolase/kinase enzyme named DLF has been constructed by gene fusion through overlap extension. This fusion enzyme consists of monomeric fructose-1,6-bisphosphate aldolase (FBPA) from *Staphylococcus carnosus* and the homodimeric dihydroxyacetone kinase (DHAK) from *Citrobacter freundii* CECT 4626 with an intervening five amino acid linker. The fusion protein was expressed soluble and retained both kinase and aldolase activities. The secondary structure of the bifuctional enzyme has been analysed by CD spectroscopy, as well as that of the parental enzymes, in order to study the effect of the covalent coupling of the two parent proteins on the structure of the fused enzyme. Since *S. carnosus* FBPA is a thermostable protein, the effect of the fusion on the thermal stability of the bifunctional enzyme has also been studied. The proximity of the active centres in the fused enzyme promotes a kinetic advantage as the 20-fold increment in the initial velocity of the overall aldol reaction indicates. Experimental evidence supports that this increase in the reaction rate can be explained in terms of substrate channelling.

Keywords: aldolases • biocatalysis • circular dichroism • fusion enzymes • protein engineering

Introduction

The aldol addition reaction has long been recognized as one of the most useful tools that has the synthetic chemist for the construction of new carbon–carbon bonds.[1] Concomitant with the C–C bond forming process is the formation of one or two new stereocenters, allowing to approach synthesis of a broad range of both natural and novel compounds. Aldolases are among the most important biocatalysts used by the Nature for C–C bond formation. Dihydroxyacetone (DHA) phosphate dependent aldolases have been thoroughly used to synthesize carbohydrates, carbohydrate-like structures, or non-carbohydrate compounds.[2] Their major synthetic advantage is that the stereochemistry of the two newly formed stereogenic centers is controlled by the enzymes and moreover, they are stereocomplementary, in that they can synthesize the four possible diastereoisomers of vicinal diols from achiral aldehyde acceptors and DHAP. However it's also well known that their major drawback is their strict specificity for DHAP. Besides the scope limitation that this fact represents, DHAP is expensive to be used stoichiometrically in high-scale synthesis and labile at neutral and basic pH, which causes that its effective concentration diminishes with time in the enzymatic reaction media. Efforts to overcome DHAP dependence of aldolases have involved *in situ* formation of arsenate or borate esters of DHA, which act as phosphate ester mimics,[3] the use of Enzyme Directed Evolution strategies to modify their donor substrate specificity,[4] development of catalytic antibodies and small peptides with aldolase activity,[5] *de novo* design of retro-aldol enzymes[6] and the use of newly discovered enzymes.[7] Besides these efforts, an efficient method of DHAP preparation is still essential and several chemical and enzymatic routes of DHAP synthesis have been described in the literature.[8] With respect to chemical synthesis, those starting from the DHA dimer[9] or 1,3-dibromoacetone,[10] provide a stable precursor of DHAP. On the other hand, enzymatic preparation of DHAP is usually coupled with the aldolic condensation catalyzed by the aldolase. Thus, DHAP can be prepared by oxidation of L-glycerol 3-phosphate (L-G3P) catalyzed by glycerophosphate oxidase, coupled with hydrogen peroxide decomposition by catalase.[11] Later on, this multi-enzyme system has been coupled with *in situ* preparation of D,L-G3P either by phosphorylation of glycerol catalyzed by the phosphatase phytase,[12] or by regioselective opening of the *rac-*

[a] Dr. L. Iturrate, Dr. I. Sánchez-Moreno, I. Oroz-Guinea, Dr. E. García-Junceda
Departamento de Química Bioorgánica
Instituto de Química Orgánica General, CSIC
Juan de la Cierva 3. 28006, Madrid, Spain
Fax: (+34) 915-644-853
E-mail: eduardo.junceda@iqog.csic.es

[b] Dr. J. Pérez-Gil
Departamento de Bioquímica y Biología Molecular, Facultad de Biología
Universidad Complutense, 28040 Madrid, Spain.

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author.

glycidol epoxide ring with phosphate.[13] A cascade reaction for the *in situ* generation of DHAP using the acid phosphatase from *Shigella flexneri* and pyrophosphate (PPi) as phosphoryl donor has also been described.[14] Direct kinase-catalysed phosphorylation of DHA using ATP as phosphoryl donor constitutes another strategy to obtain DHAP. This approach was first described in 1983 by Wong and Whitesides using the enzyme glycerol kinase.[15] Yeast ATP-dependent DHAK's have also been used for the simple and efficient preparation of DHAP.[16]

Our research group has developed a straightforward multi-enzyme system for one-pot C-C bond formation catalysed by DHAP-dependent aldolases that allows the use of DHA as initial donor. This system is based on the use of the recombinant ATP-dependent DHAK from *C. freundii* CECT 4626 for the *in situ* phosphorylation of DHA.[17] This enzyme has the highest catalytic efficiency for DHA phosphorylation described in the literature.[18] The multi-enzyme system was completed with the regeneration of ATP catalysed by acetate kinase (Scheme 1). This multi-enzyme system is attractive since it is a one-pot/one-step route to the phosphorylated aldol adduct and we have showed its utility with the three synthetically useful DHAP-dependent aldolases and with a great variety of commercially available aldehydes.[17]
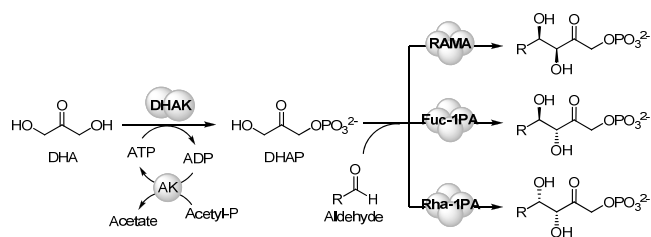
Scheme 1. Multi-enzyme system for C-C bond formation catalysed by DHAP-dependent aldolases, based in the *in situ* phosphorylation of DHA catalysed by DHAK from *C. freundii*.

However, there are at least two aspects that can be optimized:

i. The number of enzymes that take part on the system. Although the different enzymes can be overproduced in recombinant organisms, the purification of proteins is widely recognized as technically and economically challenging, accounting for a substantial fraction of the total manufacturing cost, being one of the limiting step in bioprocess development.[19]

ii. The high sensitivity of the ATP regeneration system to pH and organic solvents.[17b]

Herein, we offer an approach to solve the first point. One way to reduce the number of enzymes to express and purify is to splice two or more enzymatic activities in only one polypeptide creating a hybrid or fusion enzyme.[20] Fusion proteins have great number of applications in biotechnology, ranging from analytical procedures[21] to metabolic engineering,[22] being the purification of proteins a main application.[23] The use of fusion proteins in biotransformation may have another, even more interesting effects due to the physical association of enzymes that catalyse sequential reactions into covalently linked complexes. The close proximity of active sites of two enzymes in multi-enzyme complexes can provide a substantial substrate channelling contribution, in addition to the random

diffusion path.[24] Some of the potential catalytic advantages of substrate channelling include: i) decrease of the transit time required for an intermediate to reach the active site of the subsequent enzyme; ii) reduction of the transient time for the system to reach the new steady state; iii) protection of chemically labile intermediates; iv) circumvention of unfavorable equilibria and kinetics imposed by bulk-phase metabolite concentrations and v) segregation of the intermediates from competing chemical and enzymatic reactions.[24,25] In spite of these numerous potential advantages, fusion enzymes have been scarcely used in the biotransformation field.[26]

In the present work, we have engineered a bifunctional enzyme (named DLF), which joins the aldolase and the kinase activities in the same polypeptide.[27] We describe here the design, expression and the structural and functional characterization of this new biocatalyst.

Results and Discussion

Design of the fused aldolase/kinase enzyme: A key point at the moment of designing a fusion protein is to consider the possible interactions between subunits that can take place in the context of the chimera. When two monomeric proteins are fused, it is expected that the resultant fusion protein keeps the monomeric character of the parent enzymes. However, if the parent proteins have a dimeric or oligomeric character, the quaternary structure of the fusion protein is difficult to predict and the possible formation of macromolecular complexes of different size and activities can be obtained (Scheme 2).[20a]
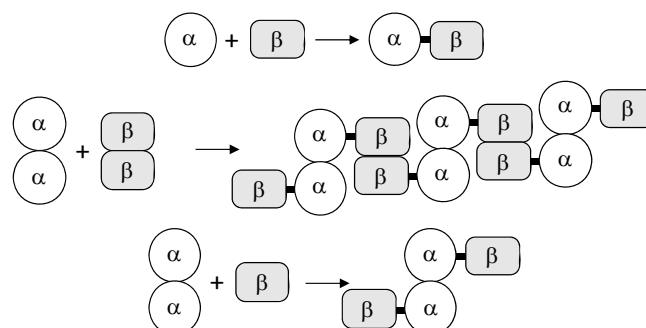
Scheme 2. Representation of the quaternary structure that can acquire a fused protein depending on the oligomeric nature of the parental proteins (adapted from ref. 20a).

Since DHAK from *C. freundii* is a dimeric enzyme,[28] we have chosen, as aldolase partner, the fructose-1,6-bisphosphate (FBP) aldolase from *Staphylococcus carnosus* because of its monomeric structure.[29] This aldolase belongs to the class I, which mechanism of reaction proceed through the formation of a Schiff-base intermediate between the donor substrate and a highly conserved lysine residue in the active site of the enzyme. In addition, this aldolase has a high pH and temperature stability.[29a] Kula and co-workers postulate that both stabilities are result of its monomeric structure and of a high renaturation potential. The crystal structure of FBPA from *S. carnosus* is unknown but the protein has a high degree of identity (55%) with the enzyme from *Porphyromonas gingivals*, which crystal structure has been resolved.[30] Figure 1 represents the 3D structure of FBPA from *S. carnosus* modeled by

2

the SWISS-MODEL server[31] using as template the structure of the enzyme from *P. gingivals*.

Figure 1. 3D structure of the FBPA from *S. carnosus* modelled by the SWISS-MODEL server. The putative catalytic lysine is showed in ball and sticks.

Another key point in the designing of a fusion protein is the selection of the linker sequence. Linker length and sequence composition can affect folding and activity of the chimeric protein. The linker sequence that has been selected to join the aldolase and the kinase enzymes was designed with the aid of the LINKER program.[32] The linker has to be long enough and sufficiently flexible to allow for the proper folding of each protein domain but short enough to keep the active sites close, in order to invigorate possible substrate channelling effects. In this sense, sequences with preference to adopt extended conformation are more desirable than sequences with high propensity to adopt α-helix or β-strand structures, which could limit the flexibility of the system. We fixed arbitrarily a five amino acids length for the linker sequence and a series of sequence filters such as the absence of polar, charged or hydrophobic residues, were introduced in the LINKER program. With this input, LINKER automatically generated a set of three peptide sequences known to adopt extended conformations (Figure 2). The sequence Gln-Gly-Gln-Gly-Gln was chosen because the DNA coding sequence does not contain restriction sites for the enzymes *Xho*I and *Nde*I chosen for the cloning of the fusion enzyme; it is a so flexible sequence that must allow the native folding of the fused enzymes and it is short enough to keep close the kinase and aldolase active centers.
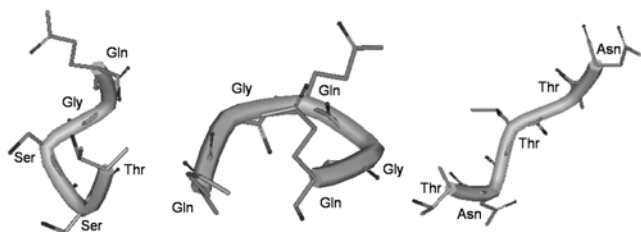

Figure 2. 3D structure of the three peptide sequences generated by the LINKER program. The structure shown is the one that these peptides present in the context of their corresponding protein.

Finally, since the C-terminal of FBPA is closer to the active center that the N-terminal end, we decided to carry out the fusion through this end. Therefore, the designed construction of the new bifunctional enzyme was DHAK-Linker-FBPA (Scheme 3).
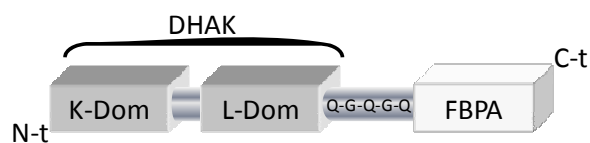
Scheme 3. Disposition of the different elements in the designed bifunctional aldolase/kinase enzyme (DLF).

Construction and characterization of the bifunctional DLF enzyme: To construct the bifunctional DLF enzyme, the genes *dhak* from *C. freundii* and *fda* from *S. carnosus* were 'spliced' together by overlap extension.[33] This method comprises two PCR steps. In the first PCR step, the *dhak* and *fda* genes were amplified separately including sites for the restriction enzymes *Nde*I (5'-*dhak*) and *Xho*I (3'-*fda*). Extremes 3'-*dhak* and 5'-*fda* included the 15 nucleotides coding for the linker sequence (*l*). These amplified genes were used as templates in a second PCR step (Figure 3).
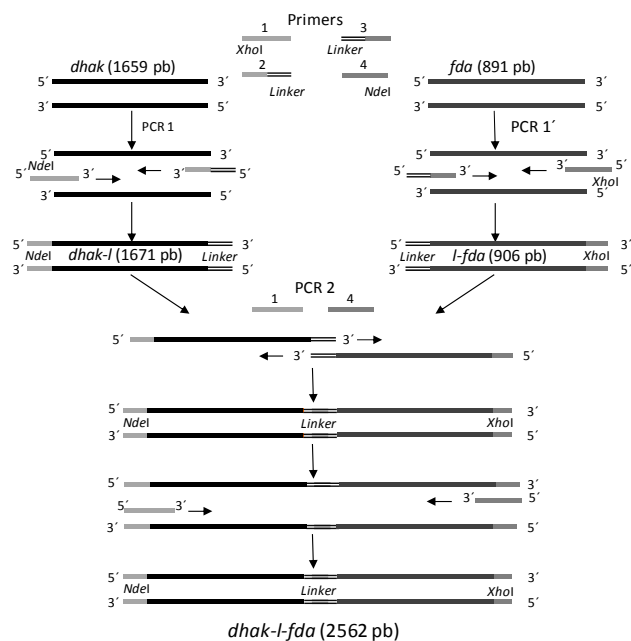

Figure 3. Strategy used for *dhak* and *fda* gene splicing by overlap extension.

The construction *dhak-l-fda* (2565 bp) was cloned into the pET-28b(+) expression vector that include a N-terminal histidine tag extending the recombinant protein. SDS-PAGE analysis of DLF expression in *E. coli* BL21(DE3) showed that DLF was expressed soluble and represented 75% of the total protein with a productivity of 135 and 425 U/L of culture broth for the kinase and aldolase activities respectively (ratio kinase:aldolase activities 1:3).

DLF was purified by immobilized metal affinity chromatography (IMAC) followed by size-exclusion chromatography. Peptide mass fingerprinting verified that the purified protein had the DLF expected features. Twenty peptides covering the DHAK sequence and 13 peptides covering the FBPA sequence were identified (Figure 4).

3

2392.1195   2203.9716
SQFFFNQRTHLVSDVIDGTIIASPWNNLARLESDPAIRIVVRRDLNKNNVAVISGGGSGHEPAHVGFIGKGMLT
1204.5004   1189.5472
AAVCGDVFASPSVDAVLTAIQAVTGEAGCLLIVKNYTGDRLNFGLAAEKARRLGYNVEMLIVGDDISLPDNKHP
1008.4979   1006.5048   1668.7333   2220.9344
926.4219   1106.5243
RGIAGTILVHKIAGYFAERGYNLATVLREAQYAANNTFSLGVALSSCHLPQEADAAPRHHPGHAELGMGIHGEP
1058.4949   2006.9221
GASVIDTQNSAQVVNLMVDKLMAALPETGRLAVMINNLGGVSVAEMAIITRELASSPLHPRIDWLIGPASLVTA
1381.5256
LDMKSFSLTAIVLEESIEKALLTEVETSNWPTPVPPREISCVPSSQRSARVEFQPSANAMVAGIVELVTTTLSD
1946.8170   938.4716
LETHLNALDAKVGDGDTGSTFAAGAREIASLLHRQQLPLDNLATLFALIGERLTVVMGGSSGVLMSIFFTAAGQ
1086.4090
KLEQGASVAESLNTGLAQMKFYGGADEGDRTMIDALQPALTSLLTQPQNLQAAFDAAQAGAERTCLSSKANAGR
2595.0438   1506.7984
ASYLSSESLLGNMDPGAHAVAMVFKALAESELGQGQGQNQEQFDKIKNGKGFIAALDQSGGSTPKALKDYGVEE
2835.1508   1176.6259   2049.0684
NEYSNDEEMFNLVHDMRTRIITSPAFNGEKILGAILFEQTMDREVEGKYTGSYLADKGIVPFLKVDKGLAEEAD
2139.1001   2085.5472
GVQLMKPIPDLDKLLDRANERGIFGTKMRSNILENNKEAIEKVVKQQFEVAKEIIAAGLVPIIEPEVNINAKDK
1501.7856
EAIEANLAEAIKAELDNLKKDQYVMLKLTIPTKVNAYSELIEHPQVIRVVALSGGYSRDEANKILKQNDGLIAS
1978.9352   1767.9388   1207.6065
FSRALVSDLNAQQSDAEFNEKLQEAIDTIFDASVNKA
1734.8908

Figure 4. Peptide mass fingerprint of bifunctional DLF. The sequence of the identified peptides are shaded and underlined. Molecular mass of each peptide is indicated in Da. Linker sequence is underlined.

Sedimentation equilibrium analysis of the purified protein confirmed the expected molecular weight and showed that the fusion protein was a homodimer, as native DHAK (Figure 5).

Figure 5. Sedimentation equilibrium analysis of FBPA (●) and DLF (○).

To analyse the effect of the covalent coupling of DHAK and FBPA in the fused enzyme, we compare its kinetic behaviour and secondary structure with those from the parent enzymes.

The kinetic parameters of DLF enzyme were calculated independently for DHA phosphorylation, to analyse the kinase activity, and for the retro-aldol reaction of fructose-1,6-bisphosphate to evaluate the aldolase activity. DLF showed a Michaelis-Menten kinetic for both activities as did the parent enzymes. The covalent union of both enzymes did not modify substantially either the K_M or the k_{cat} of DLF aldolase activity (Table 1). On the other hand, both constants were slightly modified on DLF kinase activity (Table 1). The K_M increase and the k_{cat} decrease of about three times, results in a loss of catalytic efficiency (k_{cat}/K_M) of the kinase activity in the fusion enzyme of about one order of magnitude. This loss in catalytic efficiency could be a direct result of the entropy loss of the two active sites in the bifunctional enzyme compared with the independent proteins. Apart of this, it cannot be discarded that some

active centers of the kinase had lost accessibility for the substrate in the fusion enzyme. However, despite this decrease, the k_{cat}/K_M value for this activity in the fusion enzyme, in the order of 10^6 (s^{-1}/M), is high enough to allow its use in C-C coupling reactions.

Table 1. Summary of the kinetic constants of the bifunctional DLF and of the parent enzymes.

	DHAK[a]	FBPA[b]	DLF	
			kinase	aldolase[b]
K_M (M)	1.22x10⁻⁶	1.46x10⁻⁵	3.80x10⁻⁶	1.88x10⁻⁵
k_{cat} (s⁻¹)	24.13	16.76	8.16	20.30
k_{cat}/K_M (s⁻¹/M)	1.98x10⁷	1.16x10⁴	1.19x10⁶	1.08x10⁴

[a] Constant determined for DHA phosphorylation; data taken from reference 18. [b] Constant determined in retro-aldol reaction with FBP as substrate.

Circular dichroism (CD) study of DLF secondary structure: The far-UV CD spectra obtained for the fused and the parent enzymes are indicative of a predominantly α-helical structure with two minima at ~208 and ~222 nm (Figure 6a).

Figure 6. CD spectra at 20 ºC of (a) DHAK (○), FBPA (▽) and DLF (□) and (b) equimolar mixture of DHAK and FBPA.

The contribution of each type of secondary structure is summarized in Table 2.

Table 2. Summary of the contribution of the different secondary structures components in DHAK, FBPA, DLF and DHAK:FBPA mixture.

	α-Helix[a]	β-Sheet[a]	Turns[a]	Unordered[a]
DHAK	41.4 ±0.95	13.2 ±0.93	16.8 ±0.46	28.2 ±1.95
FBPA	67.2 ±3.16	5.0 ±2.80	11.4 ±3.05	17.4 ±4.20
DLF	46.2 ±1.81	11.2 ±0.98	16.2 ±0.65	26.0 ±1.54
Mixture[b]				
Exp.[c]	52.7 ±2.50	8.3 ±0.46	15.0 ±1.33	24.7 ±1.80
Calc.[d]	53.39	10.86	15.72	25.75

[a] Contribution of each type of secondary structure is expressed in percentage. [b] Equimolar DHAK:FBPA mixture. [c] Values obtained from the CD spectra of the DHAK:FBPA mixture. [d] Values calculated applying eqn. 1 (see Experimental section) to the data obtained for the parent DHAK and FBPA enzymes.
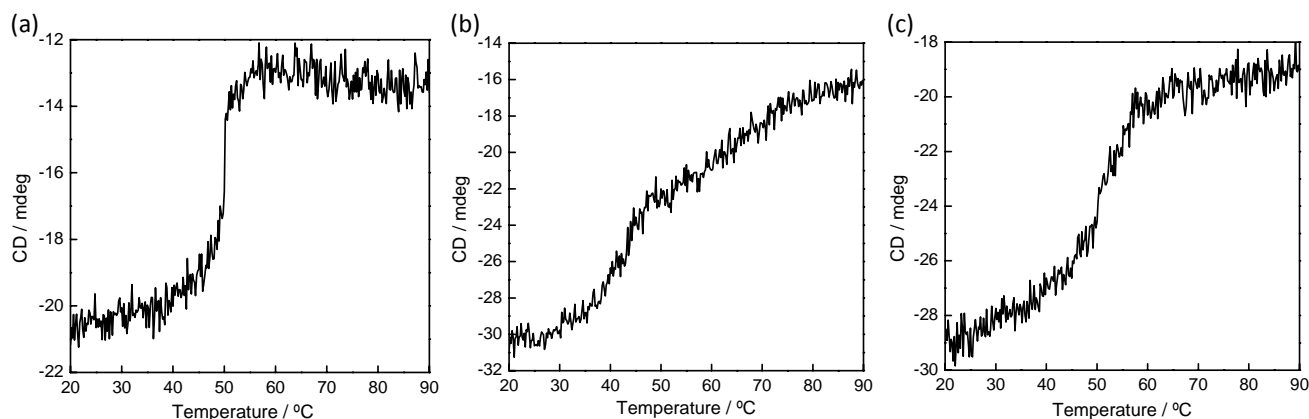
Figure 7. Variation with the temperature of the CD spectra at 222 nm of (a) DHAK, (b) FBPA and (c) DLF.

From the CD spectra of DHAK and FBPA, one can calculate the structural composition of an equimolar mixture of these enzymes (eq. 1; see Experimental Section). The calculated values should match the structural composition of DLF if the fusion has no significant effect on the folding of the fused protein domains. DLF has a contribution of α-helix (46.2 % ± 1.81) slightly lower than expected for the mixture of DHAK and FBPA (Table 2). In order to determine whether this small difference could be due to the fusion of both enzymes or to unspecific non-covalent interactions between them, the CD spectrum of an equimolar mixture of DHAK and FBPA was recorded (Figure 6b). As can be observed in Table 2, the contribution of each secondary structure component calculated from the CD spectra of the enzyme mixture, agreed with the values obtained applying equation 1. We conclude that the covalent union of both proteins is responsible for this minor reduction in α-helical conformation.

Is thermostability of FBPA transferred to the fused DLF enzyme?: As it has been commented previously, FBPA from *S. carnosus* has a considerable thermostability.[29a] Therefore, it was interesting to investigate whether this thermostability has been somehow transferred to the fusion enzyme.

Thermal stability of the three enzymes (DHAK, FBPA and DLF) was initially assessed using CD by following the temperature dependence of dichroism signal at 222 nm (Figure 7). From these curves the melting temperature (T_m) of the proteins can be calculated if they fit a two-state folding model (equation 2; see Experimental section).[34] The curve following thermal unfolding of DHAK (Figure 7a) could be fitted to a simple thermodynamic unfolding model. The melting temperature (midpoint of the transition) was determined to be 49.9 ± 0.1 °C. In addition, the transition region was very narrow and sharp, indicating that the protein existed initially as a compact, well-folded structure and that the unfolding reaction was highly cooperative. The DLF curve (Figure 7c) also responds to a two-state unfolding model, with a melting temperature of 51.8 ± 0.25 °C, significantly higher than the T_m calculated for DHAK. On the other hand, although somewhat less narrow and sharp than that of the DHAK, the width and shape of the transition region corresponds also to a initially well-folded structure and a cooperative melting process. However, the behavior of FBPA was completely different (Figure 7b). The curve clearly suggested the existence of two transitions and therefore it did not fit

to a simple thermodynamic unfolding model. Hence, we decided to analyse the thermal denaturation of FBPA in terms of "fraction of unfolding" (f_u),[35] that is defined here as the fraction of the total loss of α-helix that has occurred at each particular temperature. For comparative purposes we apply the same analysis to DHAK and DLF. As can be deduced from Figure 8, the T_m calculated for DHAK and DLF from the temperature dependence of the fraction of unfolding (48.7 °C and 53.7 °C respectively) are consistent with the T_m calculated from the melting curves in Figure 7. Figure 8 also shows that the melting temperature of the fusion enzyme DLF is almost coinciding with that of FBPA (54.9 °C).
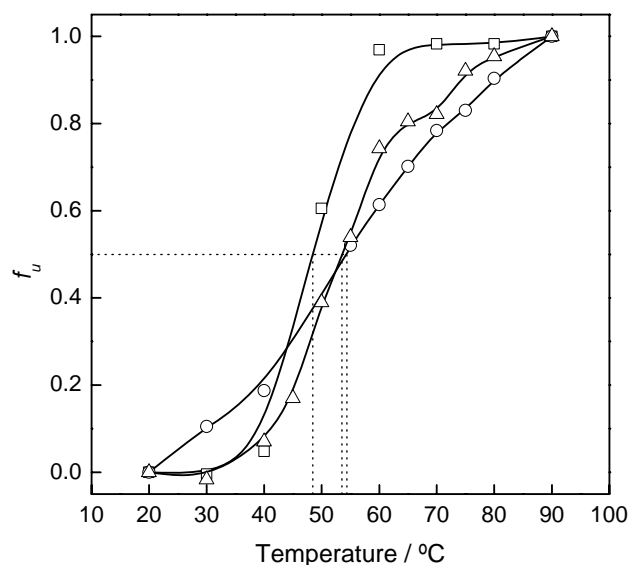
Figure 8. Thermal denaturation profiles of (□) DHAK, FBPA (○) and DLF (△).

As we did in the study of the secondary structure, to evaluate if the behavior of the fused enzyme was due to the covalent union of DHAK and FBPA or just to the simultaneous presence of both proteins, we studied the thermal denaturation of an equimolar mixture of DHAK and FBPA (Figure 9). Thermal denaturation of the mixture fit a two-state folding model with a transition region wider than those of DHAK and DLF, which is consistent with the presence of a heterogeneous population of folded structures. On this

5

occasion we found a disparity of 5 ºC between the T_m calculated from the CD spectrum at 222 nm (44.2 ± 0.5 ºC) (Figure 9a) and that calculated from the fraction of unfolding (49.3 ºC) (Figure 9b). Anyway, both values are lower than the T_m calculated for the fused enzyme (53.7 ºC). Therefore, the covalent union of FBPA and DHAK allows the thermostability of FBPA to be transferred, at least in part, to the fusion protein, making of DLF a more thermostable biocatalyst than the multi-enzyme system.
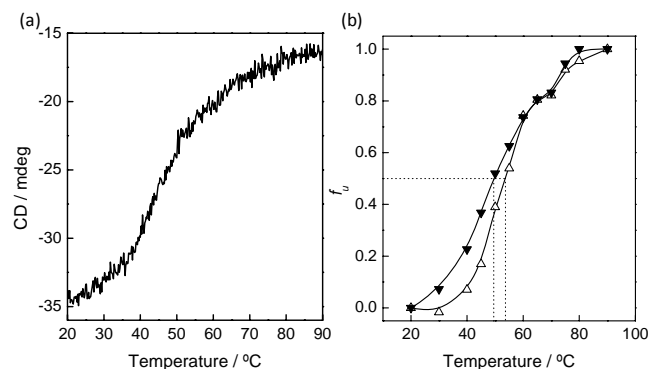

Figure 9. Thermal denaturation profiles of a equimolar mixture of DHAK and FBPA. (a) Dichroism at 222 nm, (b) variation of fraction of unfolding (▼) with temperature. The denaturation profile of DLF (△) is included for comparative purposes.

However, the T_m is directly related to conformational stability only if the unfolding is fully reversible. Irreversibility of the melting reaction is usually associated with protein precipitation or aggregation as it is unfolded. If this is the case, the melting temperature will not only depend on the intrinsic conformational stability but also on the kinetics of aggregation and the solubility of the unfolded form of the molecule. As can be observed in Figure 10, only FBPA showed an almost complete recuperation (92 %) of the secondary structure after denaturation at 90 ºC (Figure 10b). On the contrary, DHAK only recovered 51 % of its original secondary structure (Figure 10a), showing that this protein tends to precipitate and/or aggregates after denaturing. This result can be explained by the particular structure of DHAK. This protein is a homodimer and each subunit is formed by two domains clearly differentiated.[28] The K-domain is where the DHA binding site is located and the L-domain bears the ATP binding site. The ATP binding domain is a barrel formed by eight amphipathic α-helices stabilized by a phospholipid molecule (Figure 11). The loss of this lipid after unfolding could explain the tendency of DHAK to aggregate. One more time we could observe that the fusion was not neutral. Thus, the behavior of the equimolar mixture of DHAK and FBPA (Figure 10d) is practically the average between DHAK and FBPA unfolding (the mixture recovered 70 % of its original secondary structure). However, the fused enzyme DLF has a quite similar behavior to DHAK (Figure 10c) with an almost irreversible unfolding, although the percentage of recovered original structure (54 %) is slightly higher than the DHAK. This indicates that unfolding of the DHAK domain dominates the thermotropic behavior of the full chimeric construct.

Finally, to analyze if there was a correspondence between the structural stability studied by CD and the operational stability of the enzymes, we determine their functional half-life at 45 ºC (Fig. 12).
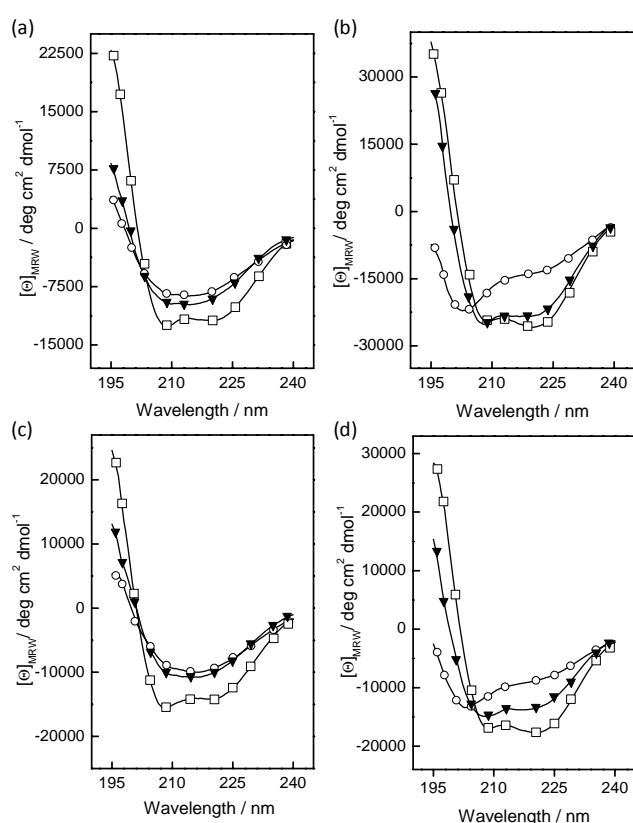
Figure 10. CD spectra at 20 ºC (□), 90 ºC (○) and again at 20 ºC after cooling the sample from 90 ºC (▼) of (a) DHAK, (b) FBPA, (c) DLF and (d) equimolar mixture of DHAK and FBPA.
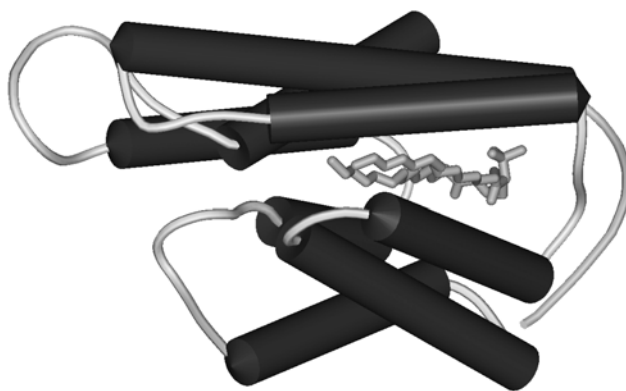
Figure 11. Crystallographic structure of the ATP binding domain from DHAK (pdb 1un9).[28] The phospholipids molecule stabilizing the folding is represented in stick.

To study how the fusion affected the activity of each enzyme, we evaluated independently the kinase and aldolase activities. DHAK, FBPA and DLF were incubated at 45 ºC over time and their remnant activities were evaluated at room temperature. We therefore determined in this assay the progressive loss of activity due to irreversible denaturation of the enzymes. The results obtained in this experiment were consistent with the results of the CD study. The half-life of the aldolase activity in the fused DLF decreased significantly with respect to the parent FBPA. This can be explained by the lower potential of renaturation of DLF with regards to FBPA. However, the half-life of the kinase activity significantly increased in the fusion enzyme with regard to the half-life of the parent

DHAK. This increment could be due to the higher T_m of the fusion enzyme.
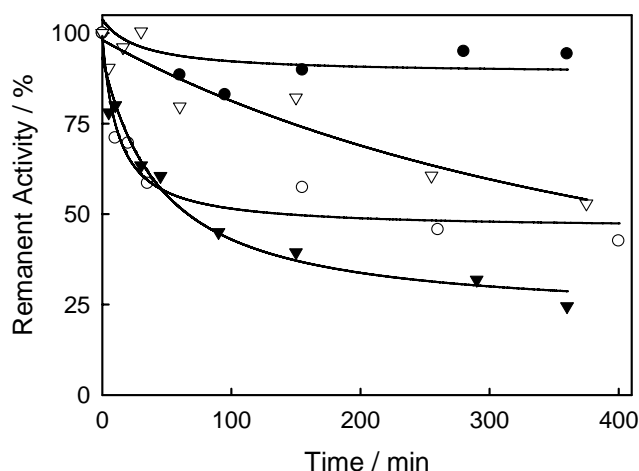

Figure 12. Operational stability at 45 ºC of (●) FBPA, (▼) DHAK, (○) aldolase activity of DLF and (▽) kinase activity of DLF.

In summary, although the aldolase activity is substantially destabilized compared to its parent enzyme, the kinase activity is markedly stabilized in the context of the fusion construct. Since the global thermal stability of both the multi-enzyme system and the fused DLF is determined by the less stable of their components, DLF should be considered a more thermostable biocatalyst than the multi-enzyme system, since the half-life of its less stable component (140 min for the aldolase activity) is more than double than the half-life of the DHAK (65 min), the least stable component of the multi-enzyme system.

Substrate channelling: As indicated above, a very interesting effect of the fusion of enzymes working sequentially is an increase of their catalytic efficiency due to the proximity of the active centres in these covalently linked multi-enzyme complexes.

Effectively, the overall rate of the coupled reaction catalysed by the fused DLF was 0.41 μmol/min whereas with the multi-enzyme system the overall rate was only of 0.02 μmol/min. That is, the physical association of the enzymes produces an increase in the aldol reaction rate of 20-folds.[27] This increase in the reaction rate could be explained in terms of substrate channelling. In the case of two free enzymes working sequentially, the product of the first reaction must diffuse through the bulk solution to meet the second enzyme. In the fusion enzyme the intermediate is directly transferred to the second enzyme without full diffusion and equilibration within the bulk solution. Therefore, DHAP produced by the kinase should have a lower transit time[36] to the aldolase active sites in the fusion enzyme compared with that of the multi-enzyme system, since the distance between the active centers has been significantly reduced. Alternatively, the obtained results can be explained from an increased local concentration of the intermediate DHAP in the proximity of the aldolase active centre producing, as a consequence, a concomitant increase in the reaction rate. In both cases free diffusion of DHAP to the bulk solution is limited.

The possible occurrence of substrate channelling can be assessed by addition to the system of a third enzyme that could trap the intermediate. The activity of this third enzyme should be higher

if the intermediate diffuse freely to the bulk solution than in the case where the diffusion of the intermediate is limited (Scheme 4).
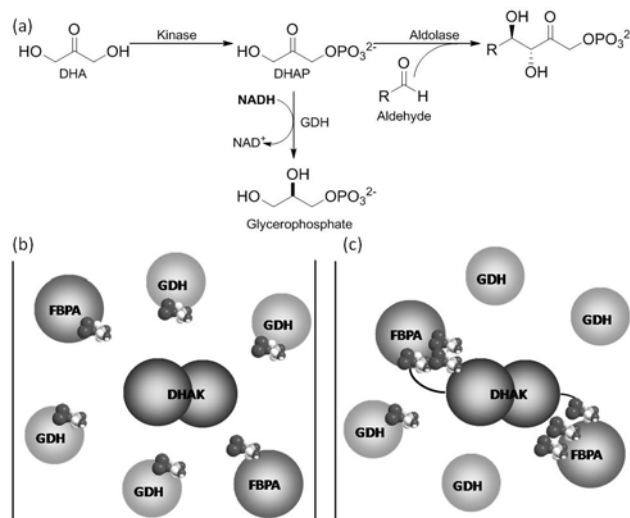
Scheme 4. Glycerol-3-phosphate dehydrogenase (GDH) catalyses the reduction of DHAP to glycerophosphate (a). In the multi-enzyme system where DHAP can freely diffuse to the bulk solution (b), the GDH activity measured must be higher than in the presence of the fusion enzyme (c) where DHAP is directly transferred to the aldolase without exposure to the bulk solution.

We carried out this experiment using glycerol-3-phosphate dehydrogenase (GDH) as trapping enzyme. GDH catalyses the reversible reduction of DHAP to glycerol-3-phosphate with concomitant oxidation of NADH (Scheme 4a). As can be observed in Figure 13, the GDH activity measured in the presence of the fused DLF enzyme or of the multi-enzyme system formed by the parental free enzymes was significantly different.
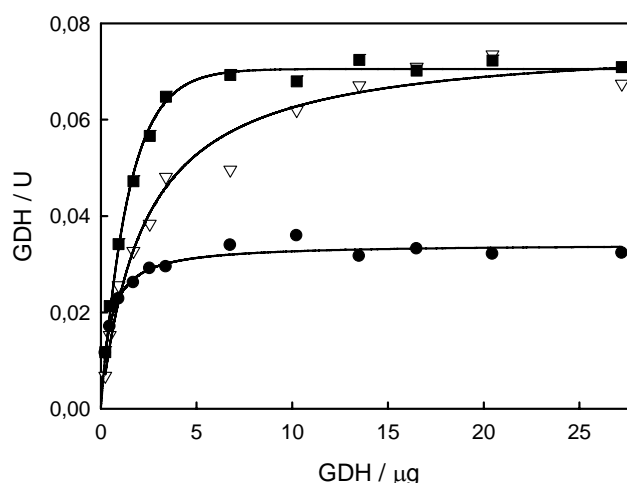

Figure 13. GDH activity measured in presence of the fused DLF enzyme (●) or in presence of the multi-enzyme system (▽). The third curve (■) represent the GDH activity measured in presence of a non functional DLF since the acceptor aldehyde was not included in the reaction mixture.

In the presence of the fusion enzyme, the maximum GDH activity measured was 0.036 U with 10 μg of enzyme. With this same amount of GDH, the activity measured in the presence of the multi-enzyme system was almost twice (0.062 U). This result

7

indicated that GDH had more substrate (DHAP) available with the multi-enzyme system than with the fusion enzyme. To confirm this interpretation, we measured GDH activity in the presence of the fusion enzyme but without adding the aldehyde acceptor. As a consequence, the second reaction catalysed by DLF—the aldol addition reaction— is unviable permitting DHAP to freely diffuse to the bulk solution. Under these conditions, the GDH activity measured reached similar values than in the presence of the multi-enzyme system (Figure 13), confirming the existence of the proposed channelling of DHAP in the fusion enzyme.

Synthetic applicability of DLF: One of the key feature to ensure that the multi-enzyme system for one-pot C-C bond formation previously described by us[17] operates optimally is that DHAP must be produced at the same rate as it is consumed by the aldolase to avoid its accumulation and to minimize its non-enzymatic degradation. Therefore, a delicate adjustment of the aldolase/kinase activity ratio was necessary. However, this adjustment is not possible when DLF is used, since in the fusion enzyme the ratio kinase:aldolase (1:3) is structurally fixed. We explored if a similar effect could be achieved modulating the ketone/aldehyde ratio. We carried out this experiment using acetaldehyde (**2**) as acceptor. In a first attempt, we used a acetaldehyde/DHA ratio of 1.5/1 equivalents. When the reaction was carried out with DLF (0.3 U of aldolase and 0.1 U of kinase), after 22 hours DHA had been almost completely consumed but only a 28% of aldol could be detected. On the other hand, there was considerable accumulation of DHAP (37%). That is, about the 35% of the produced DHAP had been degraded during the reaction without giving place to aldol formation. Doubling the amount of enzyme, we got an increase in aldol formation that reached the 44%, decreasing the accumulation of DHAP to 30.6% (Figure 14).
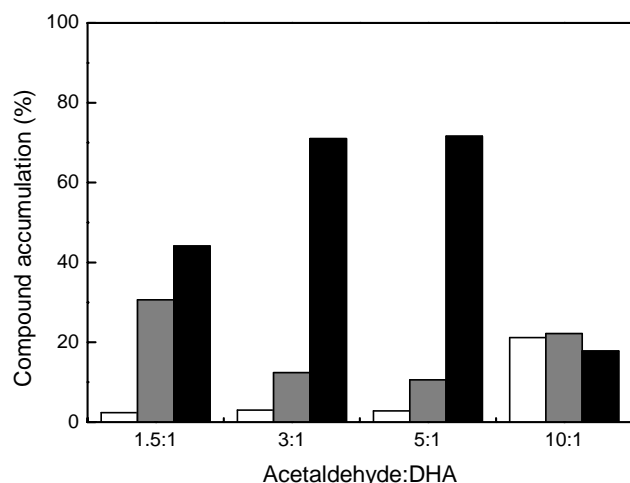
Figure 14. Figure Accumulation of DHA (□), DHAP (■) and aldol adduct (■) in the reaction mixture at different acetaldehyde:DHA ratios.

With the highest acetaldehyde:DHA ratio (10:1), it was observed that the ATP regeneration system did not work properly since after 22 h. as much as a 21% of DHA remained without react. In these conditions the aldol formation was very poor, reaching only less that 20%. On the other hand, the results obtained with the ratios 3:1 and 5: were very similar. In both cases DHA was consumed almost completely —the ATP regeneration system works correctly—, the accumulation of DHAP ranged between 10-12% and the formation of aldol adduct reached the 70% (Figure 14). That

is, modulating the aldehyde:ketone ratio we have been able to significantly increase the efficiency of the system. For subsequent studies we fixed the 3:1 ratio.

Once the optimum aldehyde:ketone ratio was determined, we studied the synthetic applicability of DLF using benzyloxyacetaldehyde (**1**), acetaldehyde (**2**) and 3-(methylthio)propionaldehyde (**3**) as acceptors (Scheme 5).[27] These reactions were conducted in small scale reactions with increased amount of DLF than in the previous experiments (1.35 U of aldolase and 0.45 U of kinase). The percentage of aldol adducts obtained after 20 h. of reaction was of 58.2 for benzyloxyacetaldehyde (**1**), 82.6 with acetaldehyde (**2**) and 67.3 with (methylthio)propionaldehyde (**3**). One more time, increasing the amount of DLF provoked an improvement in the yield of aldol formation as could be observed in the case of **5**, which yield increased from 70% to more than 80%.
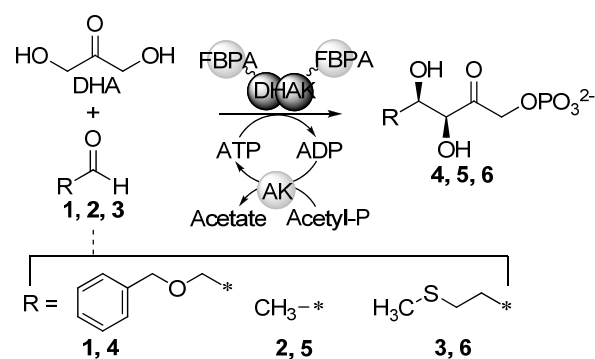
Scheme 5. Scheme of DLF application to stereoselective C-C bond formation

In a first approximation to determine if the stereoselectivity of the aldol reaction had been modified by the fusion process, we analysed the stereoisomers formed in the DLF catalysed reactions applying the enzymatic assay described by Sheldon and co-workers based on the reversibility of the aldol reaction.[37] In general, the results summarized in Table 3 are in agreement with those reported by Sheldon and co-workers for the FBPA from *S. carnosus*: the aldol reaction catalysed by DLF is highly, but far from absolutely, stereospecific for products with D-*threo* configuration and it is strongly influenced by the acceptor. In particular, the data regarding acetaldehyde (**2**) are not dissimilar to those reported for the same aldehyde by Sheldon and co-workers.[37] Therefore, we can conclude that the fusion process has not influence on the stereoselectivity of the aldol reaction.

Table 3. Stereoisomeric ratio of products formed in DLF catalysed reactions.[a]

Acceptor	Stereoisomer (%)[b]		
	D-*threo*	L-*threo*	D-*erythro*
1	79.4	2.1	18.5
2	71.8	2.0	26.2
3	96.0	0	4.0

[a] Only stereoisomers with D-*threo*, L-*threo* and D-*erythro* configuration could be detected. [b] Percentage was calculated with respect to the total aldol adduct formed.

Conclusion

As part of our ongoing project on the development of more efficient and affordable biocatalyst for C-C bond formation, we have engineered a bifunctional enzyme able to catalyse aldol addition reactions using DHA as initial ketone donor. This fusion enzyme named DLF consists of monomeric fructose-1,6-bisphosphate aldolase from *S. carnosus* and the homodimeric dihydroxyacetone kinase from *C. freundii* CECT 4626 with an intervening five aminoacids linker. The fusion protein was expressed soluble and retained both kinase and aldolase activities with a so good catalytic efficiency (k_{cat}/K_M) in the order of 10^6 (s^{-1}/M). The CD study described in this work has shown that the fusion of the parent enzymes affects only slightly the folding of the fused DLF enzyme. However, the renaturation capacity of DLF is lower than the expected for an equimolar mixture of the parent enzymes. These results could be explained because a big protein like *C. freundii* DHAK (over 120 kDa), may disturb the folding of FBPA (35 kDa), more over since FBPA is translated after DHAK.[38] The covalent union of both enzymes also allows that some of their properties are transferred to the fusion protein. Thus, the T_m of the fused enzyme is almost identical to that of FBPA and higher than would be expected if the fusion did not exercise any effect. On the other hand, we have demonstrated that the operational stability of DLF enzyme is higher than that of the multi-enzyme system. We have also shown that the physical association of the parent enzymes produces an increase in the aldol reaction rate of 20-folds that can be interpreted in terms of substrate channelling provided by the close proximity of the active centres occurred in the fusion enzyme. Finally, it seems that the fusion has not modified the steric course of the aldol reaction catalysed by the bifuctional DLF enzyme.

In summary, with the described strategy we have been able to develop a more affordable biocatalyst, more stable and more efficient than the multi-enzyme system composed by the free parent enzymes. Work is in progress in our lab to apply this strategy to other DHAP-dependent aldolases.

Experimental Section

Materials and General Procedures: *Staphylococcus carnosus* CECT 4491 was provided from the Spanish Type Culture Collection (CECT). *E. coli* BL21(DE3) competent cells were purchased from Stratagene Co. (San Diego, CA). Restriction enzymes, Taq polymerase and T4-DNA ligase were purchased from MBI Fermentas AB (Lithuania). Triosephosphate isomerase (TIM), α-glycerophosphate dehydrogenase (GDH), lysostaphin and acetate kinase (AK), were purchased from Sigma-Aldrich (St. Louis, MO). PCR primers were purchased from Isogen Life Science (Spain) and the pET-28b(+) expression vector was purchased from Novagen. Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from Applichem GmBH (Germany). Plasmids and PCR purification kits were from Promega (Madison, WI) and DNA purification kit from agarose gels was from Eppendorf (Hamburg, Germany). SDS-PAGE was performed using 10% and 5% acrylamide in the resolving and stacking gels, respectively. Gels were stained with Coomassie brilliant blue R-250 (Applichem GmBH, Germany). Electrophoresis was always run under reducing conditions, in the presence of 5% β-mercaptoethanol. Protein and DNA gels were quantified by densitometry using GeneGenius Gel Documentation and Analysis System (Syngene, U.K.). Nickel-iminodiacetic acid (Ni²⁺-IDA) agarose was supplied by Agarose Bead Technologies (Spain). Size-exclusion chromatography was carried out on a HiLoad 26/60 Superdex 75 PG column controlled using the AKTA-FPLC system (GE Healthcare Life Science). ¹H and ¹³C NMR spectra, using D_2O as solvent, were recorded on a Varian SYSTEM 500 spectrometer equipped with a 5 mm HCN cold probe with field z-gradient, operating at 500.13 and 125.76 MHz for ¹H and ¹³C, respectively. The sample temperature was maintained constant at 298 K. One-dimensional NMR experiments were performed using standard Varian pulse sequences. Two-dimensional [¹H, ¹H] NMR experiments (gCOSY) were carried out with the following parameters: a delay time of 1 s, a spectral width of 3000 Hz in both dimensions, 4096 complex points in t2 and 4 transients for each of 256 time increments, and linear prediction to 512. The data were zero-filled to 4096 × 4096 real points. Two-dimensional [¹H-¹³C] NMR experiments (gHSQC and gHMBC) used the same ¹H spectral window, a ¹³C spectral windows of 15 000 Hz, 1 s of relaxation delay, 1024 data points, and 256 time

increments, with a linear prediction to 512. The data were zero-filled to 4096 × 4096 real points. Typical numbers of transients per increment were 4 and 16, respectively.

Cloning, over-expression and purification of FBPA from *S. Carnosus*: DNA manipulation was according to standard procedures.[39] DNA template for amplification of the *fda* gene was obtained from the *S. carnosus* strain CECT 4491. The oligonucleotides 5´-ATATT<u>CATATG</u>AACCAAGAACAACAATTTGAC-3´ and 5´-TATTA<u>CTCGAG</u>TTAAGCTTTGTTTACTGA-3´ were used as leftward and rightward primers respectively (the recognition sequence for *Nde*I and *Xho*I are underlined). To extract the DNA, *S. carnosus* cells was suspended in lysis buffer containing Tris 10 mM, EDTA 10 mM, lysozyme 200 U/mL and lysostaphin 25 U/mL.[40] PCR amplification was performed in a 10 μL reaction mixture and subjected to 25 cycles of amplification. The cycle conditions were set as follows: denaturation at 94ºC for 1 min, annealing at 55 ºC for 2 min and elongation at 72 ºC for 1 min. The purified PCR product was digested with *Nde*I and *Xho*I and ligated into the doubled digested vector pET-28b(+) to yield the plasmid pET-*fda*. This plasmid was transformed into *E. coli* BL21(DE3) competent cells. A colony containing the plasmid pET-*fda* was cultured in Luria-Bertani (LB) broth containing kanamycin (26 μg/mL) at 37 ºC with shaking. When the culture reached an O.D.$_{600nm}$ of 0.5-0.6, FBPA expression was induced with IPTG (0.4 mM) and the temperature was dropped to 30 ºC. The culture was maintained over night (O/N). After that, the culture was centrifuged at 3,000 x g during 30 min at 4 ºC and the resulting pellet was treated with lysozyme and DNase for protein extraction.[41] The recombinant protein containing an N-terminal 6xHis tag was purified in a Ni⁺²-IDA-agarose column pre-equilibrated with sodium phosphate buffer (20 mM, pH 7.5). FBPA was eluted with the same buffer containing imidazole 1 M. All the fractions containing protein were pooled together and further purified by size-exclusion chromatography on a HiLoad 26/60 Superdex 75 PG column controlled using the AKTA-FPLC system (GE Healthcare Life Science). The column was developed in 50 mM phosphate buffer pH 7.2 containing NaCl (0.15 M) at a constant flow rate of 1.0 mL/min

Construction, expression and purification of the bifuctional DLF enzyme: The *dhak-l-fda* fusion gene was constructed by gene splicing through overlap extension.[33] The sequences of the primers used are shown in Table 4.

Table 4. Sequences of the primers used to splice the *dhak* and *fda* genes. The endonuclease recognition sequences are in bold and the linker sequence underlined

Fragment	Primer	Sequence
dhak-l	NtNdhak[a]	5'ATATT**CATATG**TCTCAATTCTTTTTT3'
	CtFdhak	5'<u>CTGGCCCTGGCCCTGGCCCAGCTCACTCTC</u>3'
l-fda	CtScXfda[b]	5'TATTA**CTCGAG**TTAAGCTTTGTTTACTGA3'
	NtFScfda	5'<u>CAGGGCCAGGGCCAG</u>AACCAAGAACAATTTGACAAA3'
dhak-l-fda	NtNdhak[a]	5'ATATT**CATATG**TCTCAATTCTTTTTT3'
	CtScXfda[b]	5'TATTA**CTCGAG**TTAAGCTTTGTTTACTGA3'

[a] Restriction enzyme *Nde*I. [b] Restriction enzyme *Xho*I.

The first two PCRs amplification were performed in 10 μL reaction mixture and subjected to 25 cycles of amplification. The cycle conditions were set as follows: denaturation at 94ºC for 1 min, annealing at 55 ºC for 1 min and elongation at 72 ºC for 2 min. A 1:1 mixture of the purified PCR products was used as template in the second PCR step. The reaction conditions were identical to the previously described, except for the elongation time that was 3 min. After purification, the *dhak-l-fda* fragment was digested with *Nde*I and *Xho*I and ligated into the doubled digested vector pET-28b(+) to yield the plasmid pET-*dhak-l-fda*. This plasmid was transformed into *E. coli* BL21(DE3) competent cells. Procedure for expression and purification of the bifuctional enzyme were identical to those described for the FBPA enzyme except in the IMAC purification step. In this case, previously to elute the enzyme with imidazole 1 M, the column was washed with 10 volumes of sodium phosphate buffer (20 mM, pH 7.5) containing imidazole 50 mM.

Protein analysis: Amino acid analysis of purified recombinant proteins was performed in the Protein Chemistry Service of the Centre of Biological Research (CIB-CSIC) to determine the protein concentration. The absorption spectrum of different quantified samples allowed determination of the molar extinction coefficient at 280 nm for recombinant FBPA (ε^{280} = 46292 $M^{-1}\cdot cm^{-1}$) and for DLF (ε^{280} = 77928 $M^{-1}\cdot cm^{-1}$). Peptide mass fingerprint analyses from the SDS-PAGE band corresponding to the putative FBPA and DLF were performed at the Proteomic Unit of the Spanish National Center of Biotechnology (CNB-CSIC). Samples were digested with sequencing grade trypsin O/N at 37 ºC. The analysis by MALDI-TOF mass spectrometry produces peptide mass fingerprints and the peptides observed can be collated and represented as a list of monoisotopic molecular weights. Data were collected in the *m/z* range of 800-3600. Sedimentation equilibrium experiments were performed at the Department of Chemical Physics of Biological Macromolecules (Institute of Chemical Physics

"Rocasolano"; CSIC). The initial concentration of the protein used in these experiments was 0.96 mg/mL.

Circular dichroism (CD) studies: Far-UV circular dichroism (CD) spectra were recorded in the wavelength range of 195-240 nm using a Jasco J-810 spectropolarimeter equipped with a constant temperature cell holder Jasco PTC423-S Peltier. The protein concentration was 0.2-0.4 mg/mL (3.1-6.2 μM) for DHAK, 0.1-0.2 mg/mL (2.9-5.8 μM) for FBPA and 0.6-0.7 mg/mL (6.4-7.5 μM) for DLF. The optical path length was 0.1 cm. The contribution of the buffer was always subtracted. For each sample, four spectra were accumulated at a scan speed of 20 nm/min with a bandwidth of 0.2 nm and averaged automatically. The mean residue ellipticity, $[\Theta]$, is given in units of deg·cm^2·dmol^{-1}. The secondary structure of the protein was evaluated by computer fit of the dichroism spectrum into four simple components (α-helix, β-sheet, turns and random coil) using the CDPro software package containing three commonly used programs: SELCON3, CONTIN/LL and CDSSTR (SFig. 1 and 2).[42] This software allows the use of different sets of proteins to increase the reliability of analyses. We choose the set number 4, composed of 43 soluble proteins. From data obtained for DHAK and FBPA, the theoretical structural composition of a equimolar mixture was calculated applying equation 1:

$$\psi_M = \frac{\psi_{DHAK} * PM_{DHAK} + \psi_{FBPA} * PM_{FBPA}}{PM_M} \qquad \text{(eqn 1)}$$

where ψ is each secondary structure (α-helix, β-sheet, turns or random coil) and PM is the molecular weight of the corresponding protein.

Thermal stability was assessed by following changes in the entire spectrum in the far-UV CD region with increasing temperature from 20 ºC to 90 ºC (SFig. 3). Thermal transition curves were obtained by monitoring DC signal at 222 nm as a function of temperature. The temperature gradient was 40 ºC/h and data were recorded each 0.2 ºC. These curves were fitted to a simple thermodynamic unfolding model (eqn. 2):

$$Y = (D_0 + m_D * X) - \frac{(D_0 - N_0) + (m_D - m_N) * X}{1 + \exp \frac{H_D(X - T_M)}{1.98 * (273.15 + X) * (273.15 + T_M)}} \qquad \text{(eqn 2)}$$

where D_0 is the value of the DC signal (in mdeg) of the unfolded state at T=0 ºC; m_D is the slope of the unfolded state; N_0 is the value of the DC signal (in mdeg) of the native state at T=0 ºC; m_N is the slope of the native state; T_M is the temperature (in ºC) of the inflection point and H_D is the van't Hoff enthalpy associated to the denaturation process at T=T$_M$.

The fraction of unfolding, f_u, at each temperature was calculated by dividing the amount of α-helix lost from the folded state by the total loss of α-helix (eqn. 3).

$$f_u = (h_n - h_T) / (h_n - h_u) \qquad \text{(eqn 3)}$$

where h_T is the normalized content of α-helix at a given temperature and h_n and h_u are the normalized content of α-helix at the folded and unfolded states.

Enzyme activity assays and steady-state kinetic analysis: Phosphorylation of DHA was measured spectrophotometrically in a coupled enzymatic assay as previously described.[18] Aldolase activity was spectrophotometrically measured by the retro-aldol reaction using fructose-1,6-bisphosphate (FBP) as substrate.[43] The aldolase activity assays were run at room temperature following the decrease of absorbance at 340 nm (ε_{NADH} = 6220 M^{-1}·cm^{-1}) for 5 minutes in 1 mL reaction mixture containing Tris-HCl buffer (40 mM, pH 8.0), NADH (0.2 μmol), α-GDH/TIM (2 U), 1.0 μmol of FBP and FBPA or the bifunctional DLF. Steady-state kinetic assays with DLF were measured at 25 ºC in 96-well plates in a total volume of 0.3 mL. Measurements of kinetic parameters for FBP were performed with 5.3 μg/mL of purified protein at sixteen different FBP concentrations in the range 0.005-0.25 mM. Assays to determine the kinetic parameters for DHA were performed with 24 μg/mL of purified DLF at twelve concentrations of substrate under saturating concentrations of [MgATP]$^{2-}$ complex (3.75 mM). Steady-state kinetic assays with FBPA were carried out as well at 25 ºC in 96-well plates in a total volume of 0.3 mL. Measurements of kinetic parameters for FBP were performed with 0.83 μg/mL of purified protein at twenty different FBP concentrations in the range 0.005-0.6 mM. Kinetic constants were obtained using the built-in nonlinear regression tools in SigmaPlot 8.0. For the determination of apparent kinetic constants (variation of only one substrate), initial velocities (V_i) were fitted to the Michaelis-Menten equation.

Substrate Channeling Assay: To study the possible substrate channelling promoted by the proximity of the active centres in the fused enzyme, the rates of the coupled reaction catalysed by DLF or a combination of the two parent enzymes were measured and compared under same conditions. In both cases, 0.91 U of kinase activity and 2.66 U of aldolase activity were used. To fit the Units of activity was necessary to use a slightly higher amount of DLF in terms of mg of protein. Thus, 0.49 mg of DLF was used by a total of 0.175 mg for the sum of DHAK (0.114 mg) and FBPA (0.061). The reactions were carried out at room temperature in 1.5 mL of phosphate buffer (20 mM, pH 7.5)

containing 0.05 mmol of DHA, 0.15 mmol of benzyloxyacetaldehyde (**1**), 12.5 μmol of MgSO$_4$ and 12.5 μmol of ATP. At different time, 50 μL aliquots were taken. The reaction was stopped with HClO$_4$ (7%) and the aldol product formed quantified by the retro-aldol assay. In the DHAP trapping experiment, GDH activity was measured in competition with the coupled reactions catalysed by DLF or by the multi-enzyme system. GDH activity was measured spectrophotometrically following the decrease of absorbance at 340 nm due to NADH oxidation for 10 minutes. In all cases, 0.04 U of kinase activity and 0.12 U of aldolase activity were used. The reactions were carried out at room temperature in 1 mL of phosphate buffer (20 mM, pH 7.5) containing 2.5 μmol of DHA, 7.5 μmol of acetaldehyde (**2**), 5 μmol of MgSO$_4$, 0.2 μmol of NADH, 3.75 μmol of ATP and different amounts of GDH in the range 0.231-33.99 μg.

Synthetic application of the bifunctional DLF enzyme: C-C bond formation reactions catalysed by the fusion enzyme DLF, were carried out at room temperature in 1.5 mL of phosphate buffer (20 mM, pH 7.5) containing 0.05 mmol of DHA, 0.15 mmol of aldehyde acceptor (benzyloxyacetaldehyde (**1**), acetaldehyde (**2**) and 3-(methylthio)propionaldehyde (**3**) respectively), 0.1 mmol of acetyl phosphate, 12.5 μmol of MgSO$_4$, 3 U of AK and DLF (1.5-2 U and 3-6 U of kinase and aldolase activities respectively). The reactions begin with the addition of 3.4 μmol of ATP. When the consumption of DHA was higher than 90 % (~20 h), the reactions were stopped and passed through activated carbon. The eluted was freeze-dried for NMR characterization. The determination of the stereoisomeric products formed by bifunctional DLF was carried out following the method described by Sheldon and co-workers.[37]

5-(benzyloxy)-3,4-dihydroxy-2-oxopentyl phosphate (4): *^{1}H NMR (500 MHz, D$_2$O, 298 K):* **δ** 7.2-7.1 (m, 5H, Ar), 4.52 (dd, 1H, *J*= 18.5, 5.8 Hz, H-1$_A$), 4.40 (dd, 1H, *J*= 18.5, 5.8 Hz, H-1$_B$); 4.4-4.3 (m, 2H, CH$_2$Ph), 4.29 (s, 1H, H-3), 4.15-4.05 (m, 1H, H-4), 3.50-3.45 (m, 1H, H-5$_A$), 3.45-3.40 (m, 1H, H-5$_B$). *^{13}C RMN (125 MHz, D$_2$O, 298 K):* **δ** 212.3 (C-2), 129.5 (Ar), 129.3 (Ar), 128.9 (Ar), 128.6 (Ar), 75.5 (C-1); 73.1 (C-5); 72.4 (CH$_2$Ph); 70.5 (C-4); 70.2 (C-3).

3,4-dihydroxy-2-oxopentyl phosphate (5): *^{1}H NMR (500 MHz, D$_2$O, 298 K):* **δ** 4.55 (dd, 1H, *J*= 18.8, 6.1 Hz, H-1$_A$), 4.45 (dd, 1H, *J*= 18.8, 6.1 Hz, H-1$_B$); 4.20 (d, 1H, *J*= 2.6 Hz, H-3); 4.14 (dq, 1H, *J*= 6.4, 2.4 Hz, H-4), 1.10 (d, 3H, *J*= 6.3 Hz, Me). *^{13}C NMR (125 MHz, D$_2$O, 298 K):* **δ** 211.6 (C-2), 78.5 (C-3), 67.9 (C-4), 67.7 (C-1), 18.4 (Me).

3,4-dihydroxy-6-(methylthio)-2-oxohexyl phosphate (6): *^{1}H NMR (500 MHz, D$_2$O, 298 K):* **δ** 4.55 (dd, 1H, *J*= 18.8, 6.6 Hz, H-1$_A$), 4.45 (dd, 1H, *J*= 18.8, 6.6 Hz, H-1$_B$); 4.28 (d, 1H, *J*= 2.2 Hz, H-3); 4.05 (ddd, 1H, *J*= 14.4, 11.5, 9.2 Hz, H-4), 2.5-2.4 (m, 2H, H-6), 1.95 (s, 3H, Me), 1.8-1.7 (m, 2H, H-5). *^{13}C NMR (125 MHz, D$_2$O, 298 K):* **δ** 211.5 (C-2), 77.7 (C-3), 70.4 (C-4), 68.1 (C-1), 29.6 (C-6), 23.5 (C-5); 14.3 (Me).

[1]   a) C. J. Li, *Chem. Rev.* **2005**, *105*, 3095-3166; b) *Modern Aldol Reactions*, Vol. 1 and 2 (Ed. R. Mahrwald), Wiley-VCH, Weinheim, **2004**; c) C. Palomo, M. Oiarbide, J. M. García, *Chem. Soc. Rev.* **2004**, *33*, 65-75; d) B. Alcaide, P. Almendros, *Angew. Chem.* **2003**, *115*, 884-886; *Angew. Chem. Int. Ed.* **2003**, *42*, 858-860.

[2]   For some recent reviews see: a) L. Hecquet, V. Helaine, F. Charmantray, M. Lemaire, in *Modern Biocatalysis: Stereoselective and Environmentally Friendly Reactions*, (Eds. W.-D. Fessner, T. Anthonsen), Wiley-VCH, Weinheim, **2009**, pp. 287-298; b) P. Clapes, G. A. Sprenger, J. Joglar, *ibid*, pp. 299-306; c) L. Iturrate, E. García-Junceda, in *Multi-Step Enzyme Catalysis: Biotransformations and Chemoenzymatic Synthesis*, (Ed. E. García-Junceda), Wiley-VCH, Weinheim, **2008**, 61-81; d) W.-D. Fessner, in *Asymmetric Organic Synthesis with Enzymes*, (Eds. V. Gotor, I. Alfonso, E. García-Urdiales), Wiley-VCH, Weinheim, **2008**, pp. 275-318; e) A. K. Samland, G. A. Sprenger, *Appl. Microbiol. Biotechnol.* **2006**, *71*, 253-264; f) L. J. Whalen, C.-H. Wong, *Aldrichim. Acta* **2006**, *39*, 63-71; g) J. Sukumaran, U. Hanefeld, *Chem. Soc. Rev.*, **2005**, *34*, 530-542.

[3]   a) D. G. Drueckhammer, J. R. Durrwachter, R. L. Pederson, D. C. Crans, L. Daniels, C.-H. Wong, *J. Org. Chem.* **1989**, *54*, 70-77; b) R. Schoevaart, F. Van Rantwijk, R. A. Sheldon, *J. Org. Chem.* **2001**, *66*, 4559-4562; c) M. Sugiyama, Z. Hong, L. J. Whalen, W. A. Greenberg, C.-H. Wong, *Adv. Synth. Catal.* **2006**, *348*, 2555-2559.

[4]   M. Sugiyama, Z. Hong, W. A. Greenberg, C.-H. Wong, *Bioorg. Med. Chem.* **2007**, *15*, 5905-5911.

[5] a) J. Wagner, R. A. Lerner, C. F. Barbas III, *Science*, **1995**, *270*, 1797-800; b) C. F. Barbas III, A. Heine, G. Zhong, T. Hoffmann, S. Gramatikova, R. Björnestedt, B. List, J. Anderson, E. A. Stura, I. A. Wilson,R. A. Lerner, *Science*, **1996**, *278*, 2085-2092; c) J.-L. Reymond, *J. Mol. Cat. B*, **1998**, *5*, 331-337; d) F. Tanaka, R. Fuller, C. F. Barbas III, *Biochemistry*, **2005**, *44*, 7583-7592.

[6] L. Jiang, E. A. Althoff, F. R. Clemente, L. Doyle, D. Rothlisberger, A. Zanghellini, J. L. Gallaher, J. L. Betker, F. Tanaka, C. F. Barbas III, D. Hilvert, K. N. Houk, B. L. Stoddard, Baker, D. *Science*, **2008**, *319*, 1387-1391.

[7] a) M. Schürmann, G. A. Sprenger, *J. Biol. Chem.* **2001**, *276*, 11055-11061; b) M. Schürmann, M. Schürmann, G. A. Sprenger, *J. Mol. Catal. B: Enzymatic* **2002**, *19-20*, 247-252; c) J. A. Castillo, J. Calveras, J. Casas, M. Mitjans, M. P. Vinardell, T. Parella, T. Inoue, G. A. Sprenger, J. Joglar, P. Clapés, *Org. Lett.* **2006**, *8*, 6067-6070; d) M. Sugiyama, Z. Hong, P. H. Liang, S. M. Dean, L. J. Whalen, W. A. Greenberg, C.-H. Wong, *J. Am. Chem. Soc.* **2007**, *129*, 14811-14817; e) X. Garrabou, J. A. Castillo, C. Guérard-Hélaine, T. Parella, J. Joglar, M. Lemaire, P. Clapés, *Angew. Chem.* **2009**, *121*, 5629-5633; *Angew. Chem. Int. Ed.* **2009**, *48*, 5521-5525.

[8] For an in-depth discussion of the chemical and enzymatic routes to DHAP see: M. Schümperli, R. Pellaux, S. Panke, *Appl. Microbiol. Biotechnol.* **2007**, *75*, 33-45 and references therein.

[9] a) S. H. Jung, J.-H. Jeong, P. Miller,C.-H. Wong, *J. Org. Chem.*, **1994**, *59*, 7182-7184; b) F. Charmantray, L. El Blidi, T. Gefflaut, L. Hecquet, J. Bolte, M. Lemaire, *J. Org. Chem.*, **2004**, *69*, 9310-9312.

[10] T. Gefflaut, M. Lemaire, M.-L. Valentin, J. Bolte, *J. Org. Chem.*, **1997**, *62*, 5920-5922.

[11] W. D. Fessner, G. Sinerius, *Angew. Chem.*, **1994**, *106*, 217-220; *Angew. Chem. Int. Ed.* **1994**, *33*, 209-212.

[12] R. Schoevaart, F. van Rantwijk, R. A. Sheldon, *J. Org. Chem.*, **2000**, *65*, 6940-6943.

[13] F. Charmantray, P. Dellis, S. Samreth, L. Hecquet, *Tetrahedron Lett.*, **2006**, *47*, 3261–3263.

[14] T. van Herk, A. F. Hartog, H. E. Schoemaker, R. Wever, *J. Org. Chem.* **2006**, *71*, 6244-6247.

[15] C.-H. Wong, G. M. Whitesides, *J. Org. Chem.* **1983**, *48*, 3199-3205.

[16] N. Itoh, Y. Tujibata,J. Q. Liu, *Appl. Microbiol. Biotechnol.* **1999**, *51*, 193-200.

[17] a) I. Sánchez-Moreno, J. F. García-García, A. Bastida, E. García-Junceda, *Chem. Commun.* **2004**, 1634-1635; b) I. Sánchez-Moreno, L. Iturrate, E. G. Doyagüez, J. A. Martínez, A. Fernández-Mayoralas, E. García-Junceda, *Adv. Synth. Catal.*, DOI: 10.1002/adsc.200900603.

[18] I. Sánchez-Moreno, L. Iturrate, R. Martín-Hoyos, M. L. Jimeno, M. Mena, A. Bastida, E. García-Junceda, *ChemBioChem*, **2009**, *10*, 225-229.

[19] See *J. Chem. Technol. Biotechnol.* **2008**, *83*, special issue on *Bioseparation,* (Ed. M. Rito-Palomares) and references therein.

[20] a) L. Bülow, K. Mosbach, *Trends Biotechnol.* **1991**, *9*, 226-231; b) M. Uhlén, G. Forsberg, T. Moks, M. Hartmanis, B. Nilsson, *Curr. Opin. Biotechnol.* **1992**, *3*, 363-369; c) A. E. Nixon, M. Ostermeier, S. J. Benkovic, *Trends Biotechnol.* **1998**, *16*, 258-264.

[21] T. Mizuno, K. Murao, Y. Tanabe, M. Oda, T. Tanaka, *J. Am. Chem. Soc.* **2006**, *129*,11378-11383.

[22] L. Tian, R. A. Dixon, *Planta*, **2006**, *224*, 496-507.

[23] For some reviews see: a) F. Trejo, J. L. Gelpi, M. Busquets, A. Cortes, *Curr. Top. Pept. Protein Res.* **1999**, *3*, 173-180; b) S. Stahl, S. Hober, J. Nilsson, M. Uhlen, P.-A. Hygren, *Biotechnol. Bioeng.* **2003**, *27*, 95-129; c) J. J. Lichty, J. L. Malecki, H. D. Agnew, D. J. Michelson-Horowitz, S. Tan, *Protein Expr. Purif.* **2005**, *41*; 98-105.

[24] We adopt here the operational definition of substrate chanelling given by Ovádi, according to which the term "substrate channelling" designates the coupling of two or more enzymatic reactions in which the reaction product of one enzyme is transferred to the next enzyme without escaping into the bulk phase: a) J. Ovádi, *J. Theor Biol.* **1991**, *152*, 1-22; b) H. O. Spivey, J. Ovádi, *Methods*, **1999**, *19*, 306-321.

[25] a) C. L. James, R. E. Viola, *Biochemistry*, **2002**, *41*, 3720-3725; b) C. L. James, R. E. Viola, *Biochemistry*, **2002**, *41*, 3726-3731.

[26] Some selected examples: a) P. Ljungcrantz, H. Carlsson, M. O. Mansson, P. Buckel, K. Mosbach, L. Bülow, *Biochemistry*, **1989**, *28*, 8786-8792; b) P. Ljungcrantz, L. Bülow, K. Mosbach, *FEBS Lett.* **1990**, *275*, 91-94; c) Y. Tamada, B. A. Swanson, A. Arabshahi, P. A. Frey, *Bioconjugate Chem.* **1994**, *5*, 660-665; d) X. Chen, Z. Liu, J. Wang, J. Fang, H. Fan, P. G. Wang, *J. Biol. Chem.* **2000**, *275*, 31594-31600; e) Y. H. Khang, I. W. Kim, Y. R. Hah, J. H. Hwangbo, K. K. Kang, *Biotechnol. Bioeng.* **2003**, *82*, 480-488; f) Y. Zhang, S.-Z. Li, J. Li, X. Pan, R. E. Cahoon, J. G. Jaworski, X. Wang, J. M. Jez, F. Chen, O. Yu, *J. Am. Chem. Soc.* **2006**, *128*, 13030-13031; g) L. Tian, R. A. Dixon, *Planta*, **2006**, *224*, 496-507; h) D. E. Torres Pazmiño, R. Snajdrova, B.-J. Baas, M. Ghobrial, M. D. Mihovilovic, M. W. Fraaije, *Angew. Chem.* **2008**, *120*, 2307-2310; *Angew. Chem. Int. Ed.* **2008**, *47*, 2275-2278.

[27] A preliminary account of this work has been published in communication format: L. Iturrate, I. Sanchez-Moreno, E. G. Doyaguez, E. Garcia-Junceda, *Chem. Commun.* **2009**, 1721-1723.

[28] C. Siebold, I. Arnold, L. F. García-Alles, U. Baumann, B. Erni, *J. Biol. Chem.* **2003**, *278*, 48236-48244.

[29] a) H. P. Brockamp, M. R. Kula, *Appl. Microbiol. Biotechnol.* **1990**, *34*, 287-291; b) M. T. Zannetti, C. Walter, M. Knorst, W.-D. Fessner, *Chem. Eur. J.* **1999**, *5*; 1882-1890; c) M. Dinkelbach, M. Hodenius, A. Steigel, M. R. Kula, *Biocatal. Biotransform.* **2001**, *19*, 51-68.

[30] a) C. Witke, F. Götz, *J. Bacteriol.*, **1992**, *175*,7495-7499; b) Y. Kim, S. Moy, A. Joachimiak, PDB: 2IQT.

[31] a) M. C. Peitsch, *Bio/Technology*, **1995**, *13*, 658-660; b) K. Arnold, L. Bordoli, J. Kopp, T. Schwede, *Bioinformatics*, **2006**, *22*, 195-201; c) F. Kiefer, K. Arnold, M. Künzli, L. Bordoli, T. Schwede, *Nucleic Acids Res.*, **2009**, *37*, D387-D392.

[32] a) C. J. Crasto, J.-A. Feng, *Protein Eng.*, **2000**, *13*, 309-312; b) F. Xue, Z. Gu, J.-A. Feng, *Nucleic Acids Res.*, **2004**, *32*, W562-W565.

[33] R. M. Horton, H. D. Hunt, S. N. Ho, J. K. Pullen, L. R. Pease, *Gene*, **1989**, *77*, 61-68.

[34] L. M. Mayr, O. Landt, U. Hahn, F. X. Schmidt, *J. Mol. Biol.* **1993**, *231*, 897-912.

[35] a) K.-P. Wong, C, Tanford, *J. Biol. Chem.* **1973**, *248*, 8518-8523; b) L. Z. Wu, B. L. Ma, D. W. Zou, Z. X. Tie, J. Wang, W. Wang, *J. Mol. Struct.* **2008**, *877*, 44-49.

[36] Transit time is the time required for a metabolite to reach the next enzyme in the pathway and is expected to be in the order of r^2_E/D_s, where r_E is the average separation distance of enzyme molecules and D_s is the diffusion coefficient of the substrate: G. R. Welch, *Prog. Biophys. Mol. Biol.* **1977**, *32*, 103-191.

[37] R. Schoevaart, F. v. Rantwijk, R. A. Sheldon, *Biotechnol. Bioeng.* **2000**, *70*, 349-352.

[38] A. Stanislawska-Sachadyn, P. Sachadyn, K. Ihle, C. Sydorczuk, K. Wiejacha, J. Kur, *J. Biotechnol.* **2006**, *121*, 134-143.

[39] J. Sambrook, E. F. Fritsch, T. Maniatis *Molecular cloning. A laboratory manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y, 2nd edn., **1989**.

[40] M. Ligozzi, R. Fontana, *Afr J Biotechnol* **2003**, *2*, 251-253.

[41] A. Bastida, A. Fernández-Mayoralas, R. Gómez, F. Iradier, J. C. Carretero, E. García-Junceda, *Chem. Eur. J.* **2001**, *7*, 2390-2397.

[42] N.Sreerama, R. W. Woody, *Anal. Biochem.* **2000**, *287*, 252-260.

[43] H. U. Bergmeyer *Methods of Enzymatic Analysis*, vol. 2, 3rd ed, Verlag Chemie: Deerfield, FL, **1984**.

Entry for the Table of Contents (Please choose one layout only)

Layout 1:

Engineered bifuntional aldolase/kinase or when one plus one is more than two: A bifunctional enzyme able to catalyse aldol addition reactions using DHA as initial ketone donor has been engineered. This new enzyme is a more affordable biocatalyst, more stable and more efficient than the multi-enzyme system composed by the free parent enzymes.

Preparation and characterization of a bifunctional aldolase/kinase enzyme. A more efficient biocatalyst for C-C bond formation

Laura Iturrate, Israel Sánchez-Moreno, Isabel Oroz, Jesús Pérez-Gil and Eduardo García-Junceda*

Supplementary Information
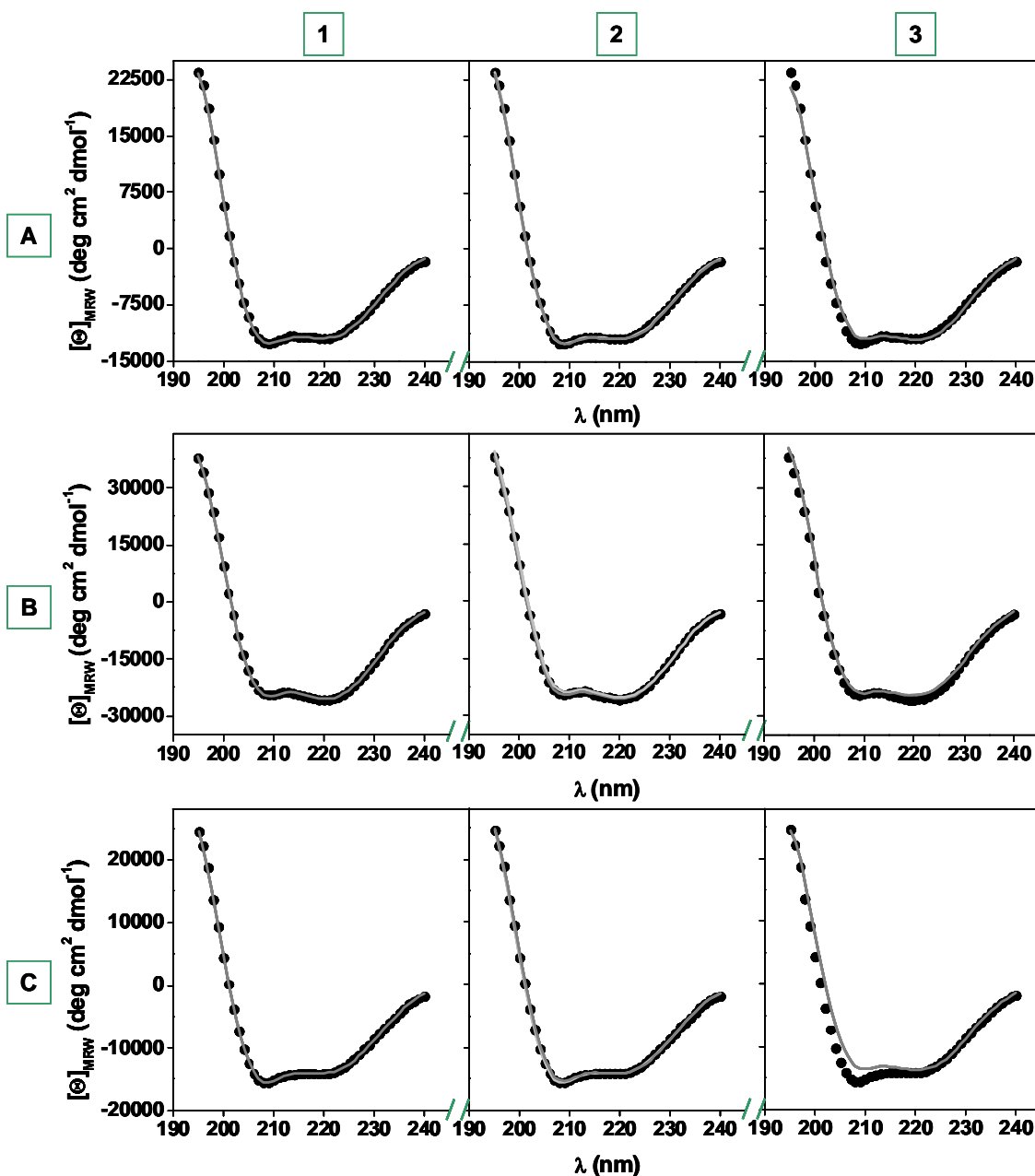
SFigura 1. CD spectra at 20 ºC of DHAK (A); FBPA (B) and DLF (C). The dots represent the experimental data and the lines show the theoretical spectra calculated with: (1) CONTIN, (2) SELCON3 (3) CDSSTR
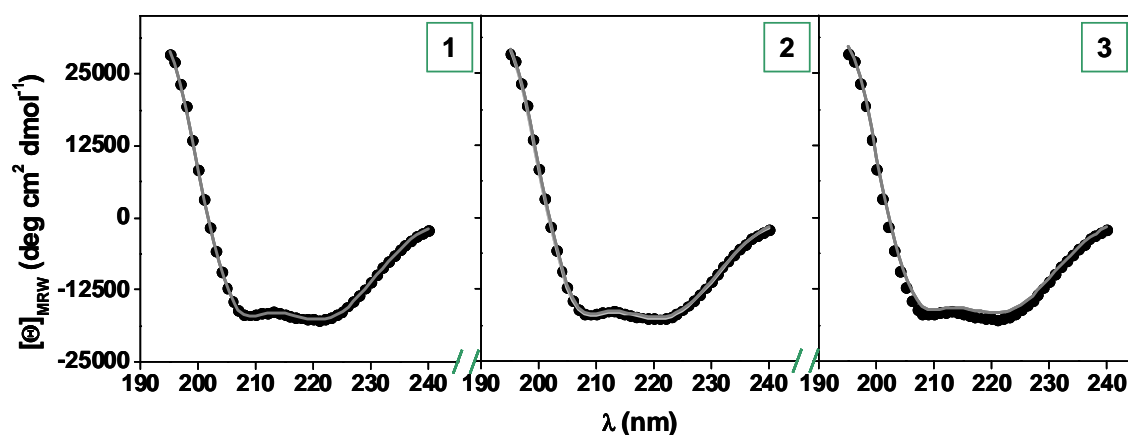
SFigura 2. CD spectra at 20 ℃ of a DHAK:FBPA mixture (mol:mol). The dots represent the experimental data and the lines show the theoretical spectra calculated with: (1) CONTIN, (2) SELCON3 (3) CDSSTR
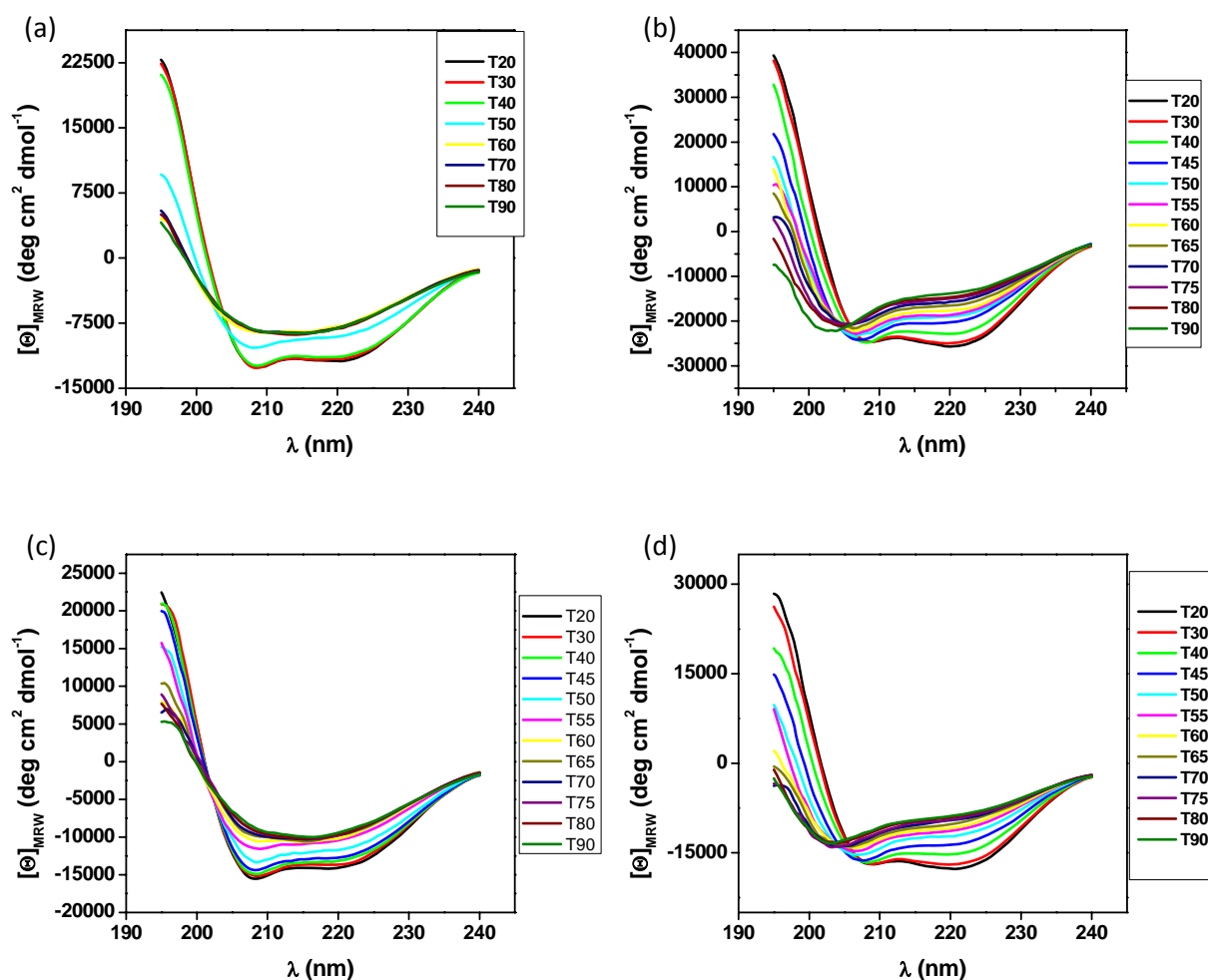

SFigura 3. CD spectra at different temperatures of (a) DHAK, (b) FBPA, (c) DLF and (d) equimolar mixture of DHAK and FBPA used to calculate the fraction of unfolding f_u.