Substrate Channelling in an Engineered Bifunctional Aldolase/Kinase Enzyme Confers Catalytic Advantage for C-C Bond Formation.

*Laura Iturrate, Israel Sánchez-Moreno, Elisa G. Doyagüez and Eduardo García-Junceda**

Supplementary Information

Contents

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 (αGDH), 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 (Ni^{2+}-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). All other chemicals were purchased from commercial sources as reagent grade.

Cloning, overexpression and purification of FBPA from S. carnosus

DNA manipulation was according to standard procedures.[1] DNA template for amplification of the *fda* gene was obtained from the *S. carnosus* strain CECT 4491. The oligonucleotides 5′-ATATTCATATGAACCAAGAACAACAATTTGAC-3′ and 5′-TATTACTCGAGTTAAGCTTTGTTTACTGA-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.[2] 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 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.[3]

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

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

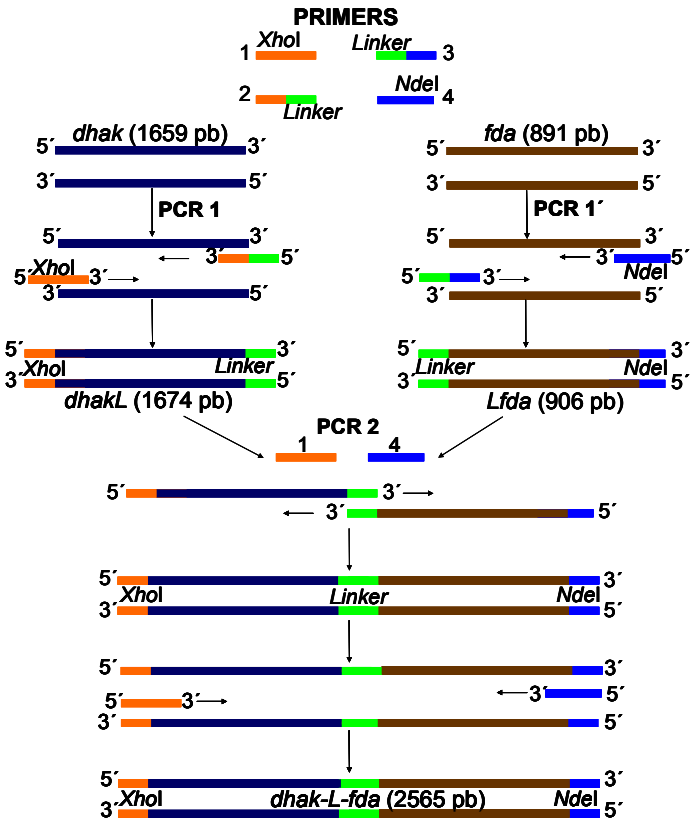
[2] M. Ligozzi, R. Fontana, *Afr J Biotechnol* **2003**, *2*, 251-253.

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

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 by overlap extension (SFigure 1).[4] This method comprises two PCR steps. In the first PCR step, the *dhak* gene from *C. freundii* and the *fda* gene from *S. carnosus* 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 linker sequence (*l*). These amplified genes were used as templates for a second PCR step.



SFigure 1. Schematic representation of DHAK and FBPA genes splicing by overlap extension.

The sequences of the primers used are shown in the STable 1. 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.

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

STable 1. 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	Endonuclease
dhak-l	NtNdhak	5'-ATATT**CATATG**TCTCAATTCTTTTTT-3'	*Nde*I
	CtFdhak	5'-CTGGCCCTGGCCCTGGCCCAGCTCACTCTC-3'	-
l-fda	CtScXfda	5'-TATTA**CTCGAG**TTAAGCTTTGTTTACTGA-3'	*Xho*I
	NtFScfda	5'-CAGGGCCAGGGCCAGAACCAAGAACAATTTGACAAA-3'	-
dhak-l-fda	NtNdhak	5'-ATATT**CATATG**TCTCAATTCTTTTTT-3'	*Nde*I
	CtScXfda	5'-TATTA**CTCGAG**TTAAGCTTTGTTTACTGA-3'	*Xho*I

Protein analysis

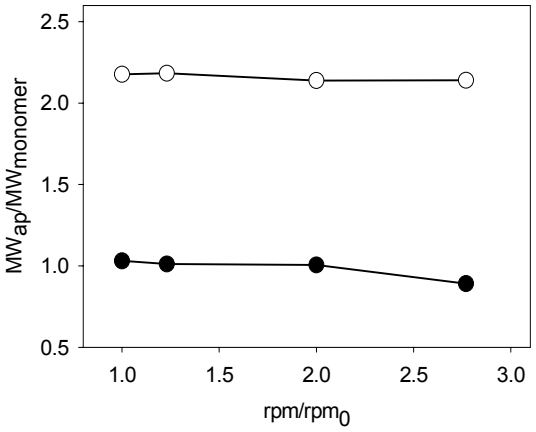
Amino acid analyses of purified recombinant proteins were performed in the Protein Chemistry Service of the Centre of Biological Research (CIB-CSIC) to determine the protein concentration. The absorption spectrum of different quantified samples allowed determination of the molar extinction coefficient 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 (SFigure 2) 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 overnight at 37ºC. 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. 20 peptides covering the major part of the amino acid sequence of FBPA were identified (SFigure 2A). Almost all the predicted tryptic peptides with molecular masses falling in the analyzed *m/z* range were found in the peptide mass fingerprint of the recombinant FBPA.

SFigure 2. Peptide mass fingerprint of FBPA from *S. carnosus* (A) and bifunctional DLF (B). The sequence of the identified peptides are shaded and underlined. Molecular mass of each peptide is indicated in Da. Linker sequence in the DLF enzyme is marked in yellow.

In the case of the DLF peptide mass fingerprinting, 20 peptides covering the DHAK sequence were identified and 13 peptides covering the FBPA sequence were also identified (SFigure 2B). Peptide mass fingerprinting verified that purified proteins had the expected features of FBPA and DLF respectively.

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. These experiments confirmed the expected molecular weight and showed that the fusion protein was a homodimer, as the native DHAK (SFigure 3).



SFigure 3. Sedimentation equilibrium analysis of FBPA (●) and DLF (○).

Enzyme activity assays

Phosphorylation of DHA was measured spectrophotometrically in a coupled enzymatic assay as previously described.[5] Aldolase activity was spectrophotometrically measured by the retro-aldol reaction using fructose-1,6-bisphosphate (FBP) as substrate.[6] The aldolase activity assays were run at room temperature following the decrease of absorbance at 340 nm (ε_{NADH}^{340} = 6220 $M^{-1}\cdot cm^{-1}$) 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 the bifunctional DLF or FBPA.

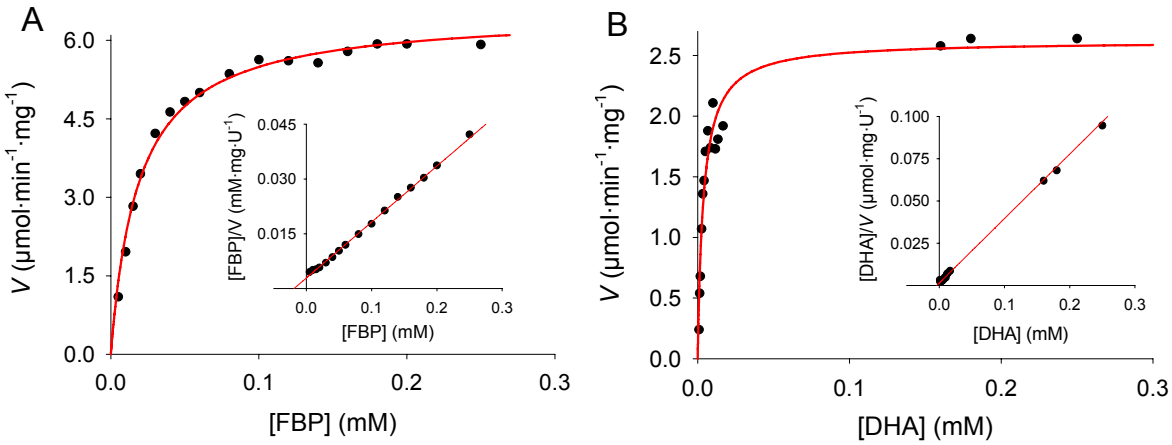
Steady-state kinetic assays

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 (SFigure 4A) 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 (SFigure 4B) were performed with 24 μg/mL of purified DLF at twelve concentrations of substrate under saturating concentrations of $[MgATP]^{2-}$ complex (3.75 mM).

The catalytic constant (k_{cat}) is the result of dividing the V_{max} (in units of M x $time^{-1}$) by the concentration (Molar) of enzyme.

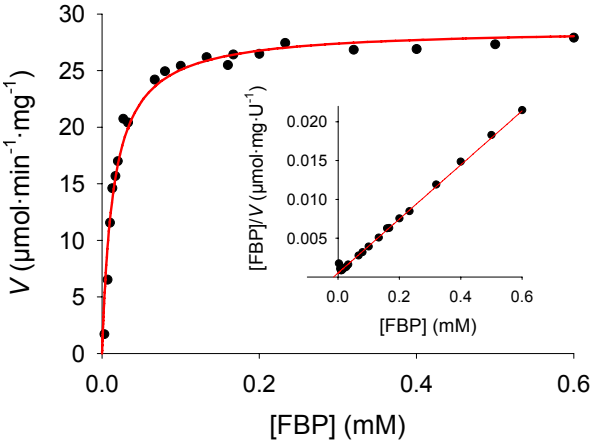
[5] I. Sánchez-Moreno, L. Iturrate, R. Martín-Hoyos, M. L. Jimeno, M. Mena, A. Bastida, E. García-Junceda, *ChemBioChem*, DOI:10.1002/cbic.200800573.

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



SFigure 4. Substrate kinetics of aldolase (A) and kinase (B) activities for fusion enzyme DLF. The inserts show Hanes-Woolf plots used for kinetic constants determination.

Steady-state kinetic assays with FBPA were measured as well at 25 ºC in 96-well plates in a total volume of 0.3 mL. Measurements of kinetic parameters for FBP (SFigure 5) 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



SFigure 5. Substrate kinetic for FBPA from *S. carnosus*. The insert shows Hanes-Woolf plot used for kinetic constants determination.

Proximity Effect study

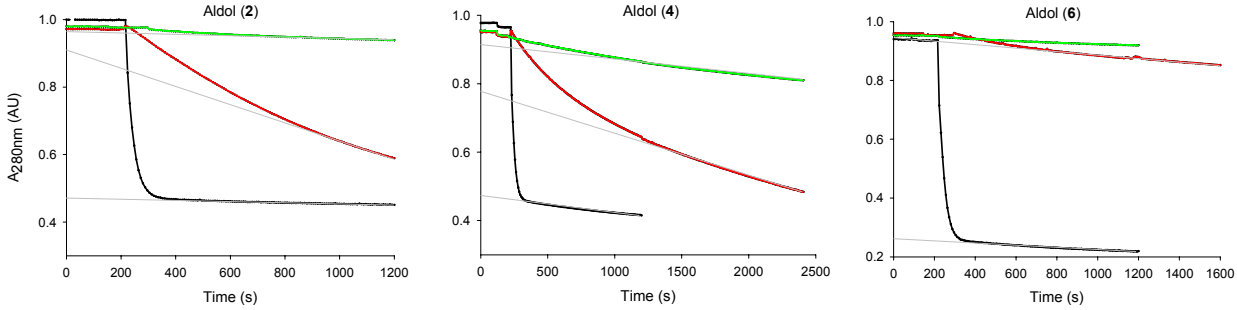
To study the proximity effect, the rates of the coupled reaction catalysed by DLF or a combination of the two native enzymes were measured and compared under same conditions. In both cases, 0.91 U of kinase activity and 2.66 U of aldolase activity were used. To fit the Units of activity was necessary 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.

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 (**3**) and 3-(methylthio)propionaldehyde (**5**) 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.

Stereochemistry study of aldol reactions catalyzed by DLF enzyme

The determination of the stereoisomeric products formed by bifuctional DLF was carried out following the method described by Sheldon and co-workers (SFigure 6).[7]

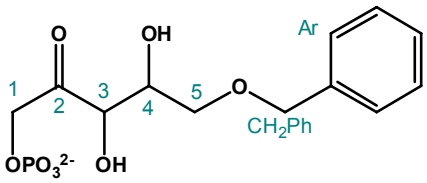


SFigure 6. Enzymatic determination of stereoisomers of aldol products (**2**), (**4**) and (**6**). Plots show retro-aldol reactions catalysed by RAMA (black line), Fuc-1PA (red line) and Rham-1PA (green line).

NMR procedures.

^1H and ^{13}C 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 ^{13}C, 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-^{13}C] NMR experiments (gHSQC and gHMBC) used the same ^1H spectral window, a ^{13}C 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.

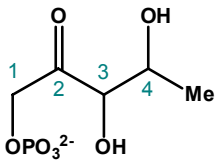
5-(benzyloxy)-3,4-dihydroxy-2-oxopentyl phosphate (**2**)



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

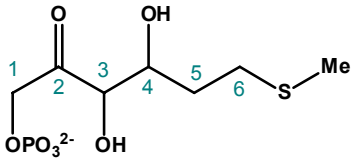
^{1}H NMR (500 MHz, D_2O, 298 K): **δ** 7.2-7.1 (m, 5H, Ar), 4.52 (dd, 1H, *J*= 18.5, 5.8 Hz, H-1_A), 4.40 (dd, 1H, *J*= 18.5, 5.8 Hz, H-1_B); 4.4-4.3 (m, 2H, CH_2Ph), 4.29 (s, 1H, H-3), 4.15-4.05 (m, 1H, H-4), 3.50-3.45 (m, 1H, H-5_A), 3.45-3.40 (m, 1H, H-5_B).

3,4-dihydroxy-2-oxopentyl phosphate (**4**)



^{1}H NMR (500 MHz, D_2O, 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_2O, 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_2O, 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_2O, 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).