Supplementary Information

Structure-Guided redesign of D-Fructose-6-Phosphate Aldolase from *E. coli*: remarkable activity and selectivity towards acceptor substrates by two-point mutation.

Mariana Gutierrez,[a] Teodor Parella,[b] Jesús Joglar,[a] Jordi Bujons[a] and Pere Clapés*[a]

[a]Instituto de Química Avanzada de Cataluña-CSIC. Jordi Girona 18-26, 08034 Barcelona, Spain. Fax:34 93 2045904; Tel: 34 93 4006112; E-mail: pere.clapes@iqac.csic.es

[b] Servei de Ressonància Magnètica Nuclear, Dept of Chemistry Universitat Autònoma de Barcelona, Bellaterra, Spain.

A) Experimental

General. All the DNA manipulations and bacterial transformation were carried out according to the standard protocols or manufacturers' instructions. Synthetic oligonucleotides were purchased from MWG-Biotech. Phosphoglucose isomerase from baker´s yeast (*S. cerevisiae*), glucose 6-phosphate deshydrogenase, DL-glyceraldehyde 3-phosphate (G3P), *N*-benzyloxyacetaldehyde and nicotinamide adenine dinucleotide phosphate ($NADP^+$) were from Sigma-Aldrich. Antibiotics, acrylamide-bisacrylamide, buffer components and hydroxyacetone (HA) were from Sigma-Aldrich. Culture media components for bacteria were from Pronadisa (Madrid, Spain). Dyhydroxyacetone (DHA) was purchased from Merck. Deionized water was used for preparative HPLC and Milli-Q-grade water for analytical HPLC. All other solvents used were of analytical grade. *N*-Cbz-aminoaldehydes used in these studies were synthesized in our lab using procedures published in previous works.[1-3] Bacterial strains, oligonucleotides and plasmid used in this study are listed in Table S1. The plasmid pQE*fsa* containing the gene for expression of FSA was constructed in our lab using routine procedures of molecular biology.

NMR analysis: High field ^1H and ^{13}C nuclear magnetic resonance (NMR) analyses were carried out using an AVANCE 500 BRUKER spectrometer equipped with a high-sensitive CryoProbe for [D2]H_2O and [D4]MeOH solutions. Full characterization of the described compounds was performed using typical gradient-enhanced 2D experiments: COSY, NOESY, HSQC and HMBC, recorded under routine conditions. When possible, NOE data were obtained from selective 1D NOESY versions using a single pulsed-field-gradient echo as a selective excitation method and a mixing time of 500 ms. When necessary, proton and NOESY experiments were recorded at different temperatures in order to study the different behavior of the exchange phenomena to avoid the presence of false NOE cross-peaks that complicates both structural and dynamic studies. Routine, ^1H (400-500 MHz) and ^{13}C (101 MHz) NMR spectra of compounds were recorded with a Varian Mercury-400 and Varian Anova-500 spectrometers, respectively.

Specific rotations. Analytical specific rotation values were measured with a Perkin Elmer Model 341 (Überlingen, Germany) polarimeter.

Protein concentration was calculated with the method of Bradford.[4]

Activity assays. The FSA activity was measured using the formation of D-fructose-6-phosphate from dihydroxyacetone (300 mM) and D,L-glyceraldehyde 3-phosphate (2.8 mM) in 50 mM glycylgycine buffer (pH 8.5) containing 1 mM dithiothreitol at 30 °C as described previously using a coupled assay using phosphoglucose isomerase and glucose-

6-phosphate dehydrogenase.[5,6] The reduction of NADP (0.5 mM) was monitored at 340 nm for 5 min using a Cary 50 Bio UV-Visible Spectrophotometer (Varian, Darmstadt, Germany); 1 mmol of NADP reduced was set equivalent to 1 mmol of D-fructose-6-phosphate formed (one U will produce 1 μmol of D-fructose-6-phosphate (D-F6P) per minute).

HPLC analyses. HPLC analyses were performed on an X-Bridge™ C18, 5μm, 4.6 x 250 mm column from Waters (Milford, USA). Samples (30 μL) were injected and eluted with the following conditions: solvent system: (A): aqueous trifluoroacetic acid (TFA; 0.1% (*v/v*)) and solvent (B): TFA (0.095% (*v/v*)) in CH_3CN/H_2O (4:1), gradient elution from 10-70 % B in 30 min, flow rate 1mL min^{-1}, detection 215nm, column temperature 30 ºC. The amount of aldol adduct produced was quantified from the peak areas using an external standard methodology. The detection limit for the Cbz-aldol adducts was between 0.2 to 0.5 nmol.

Preparative column chromatography. Bulk stationary phase Amberlite™ XAD™16 (Rohm and Haas) was packed into a glass column (45 x 1.5 cm) to give a final bed volume of 200 mL. The Amberlite™ XAD™16 was equilibrated initially with H_2O and then the crude material dissolved in water (50-70 mL) was loaded onto the column. Impurities were washed away with H_2O (2–3 column volumes). Then, the aldol adducts were eluted with EtOH (200 mL). The fractions were analyzed by HPLC. Pure fractions were pooled and lyophilized.

Mutagenesis. FSA gene mutations were introduced with the QuikChange site-directed mutagenesis kit (QuickChange® Site-Directed Mutagenesis kit, Stratagene), using the plasmid pQE*fsa* as template and performed according to the manufacturer´s protocols. The oligonucleotides for amplification and mutagenesis were custom-synthesized (Table S1). The double mutant FSA A129S/A165G was obtained using pQE*fsaFSA129S* as template with the mutagenesis primers FSA A165G up and FSA A165G down (Table S1). The external primers utilized were pQE53 and pQE35. Strain *E. coli* JM109 was used for transformations and plasmid preparation. The resulting clones were checked by DNA sequencing and the proteins expressed analyzed by mass spectra.

Table S1: Strain, plasmids and oligonucleotides used in this study (mutagenised codons are highlighted in italic and bold).

Plasmids	Relevant genetic characteristics	Ref./Origin
pQE40	P_{T5}, *Amp*^{r}, *ColE1*$_{ori}$	Quiagen
pQE*fsa*	*fsa* gene (660bp) cloned in pQE40 (*BseRI/HindIII*), the C-terminal His-tagged was deleted.	This study
pQE*fsaA129S*	*fsaA129S* gene cloned in pQE40	This study
pQE*fsaR134A*	*fsaR134A* gene (660bp) cloned in pQE40	This study
pQE*fsaA165G*	*fsaA165G* gene (660bp) cloned in pQE40	This study
pQE*fsaA129S/A165G*	*fsaA129S/A165G* gene (660bp) cloned in pQE40	This study
Strains	**Relevant genotype**	**Ref./Origin**
E. coli JM109	*end*A1, *rec*A1, *gyr*A96, *thi, hsd*R17 (r_k^-, m_k^+), *rel*A1, *sup*E44, Δ(*lac-pro*AB), [F´ *tra*D36, *pro*AB, *laq*I^{q}ZΔM15].	Promega
E. coli M15[pREP4]	*na*I^s, *str*^{s}, *rif*^{s}, *thi*$^-$, *lac*$^-$, *ara*$^+$, *gal*$^+$, *mtl*$^-$, F$^-$, *rec*A$^+$, *uvr*$^+$, *lon*$^+$.	Qiagen
Oligonucleotides	**Oligonucleotide sequences (5´→ 3´)**	**Ref./Origin**
FSAA129S up	GGCAGGTGCGGAATATGTT***TCG***CCTTACGTTAATC	This study
FSAA129S down	GATTAACGTAAGG***CGA***AACATATTCCGCACCTGCC	This study
FSAR134A up	CCTTACGTTAAT***GCT***ATTGATGCTCAG	This study
FSAR134A down	CTGAGCATCAAT***AGC***ATTAACGTAAGG	This study
FSAA165G up	GAAAGTGCTGGCA***GGG***AGTTTCAAAACC	This study
FSAA165G down	GGTTTTGAAACT***CCC***TGCCAGCACTTTC	This study

Protein expression and purification. The plasmids were transformed into *E. coli* strain M-15 [pREP-4] (QIAGEN). Cells were grown at 37ºC in 4.8 L of a LB medium containing ampicillin (100 mg L^{-1}) and kanamycin (25 mg L^{-1}) up to an optical density of 0.6 at 600 nm. For protein expression, temperature was lowered to 30ºC to avoid inclusion bodies formation and isopropyl-*β*-D-thiogalactoside (IPTG) was added (final concentration of 1 mM). After additional 4 h cells were harvested, suspended in starting buffer (50 mM glycylglycine, 1 mM dithiothreithol (DTT), pH 8.5) and lysed using a TS 0.75kW 40K Cell Disrupter (Constant Systems). Cellular debris was removed by centrifugation at 12000 *g* for 20 min. The clear supernatant was heat-shock treated (70 ºC, 30 min) and

centrifuged at 12000 *g* for 20 min at 4 ºC. The pellet was discarded and the supernatant was dialyzed against buffer glycylglycine 5 mM, DTT 0.1 mM, pH 8.5 and finally, lyophilized. Proteins were analyzed by SDS-PAGE (Figure 1).

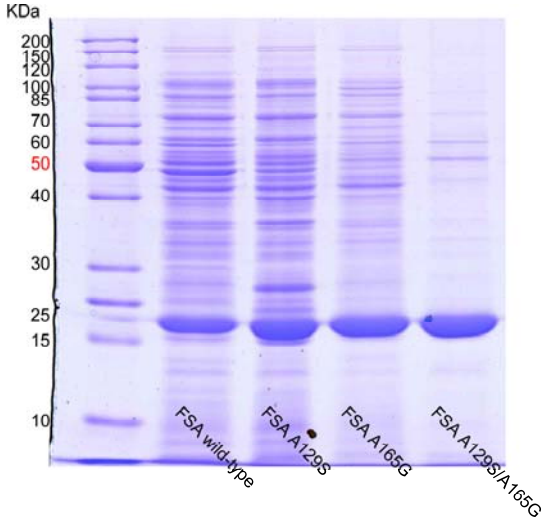


Figure S1. SDS-PAGE analysis of FSA wild-type, FSA A129S, FSA A165G and FSA A129S/A165G. Protein (5 μg in all cases; FSA wild-type, 0.75 mg of protein/g lyophilized powder; FSA A129S, 0.8 mg of protein/g lyophilized powder; FSA A165G, 0.33 mg of protein/g lyophilized powder and FSA A129S/A165G, 0.4 mg of protein/g lyophilized powder) were applied in each *lane* to a 12% (v/v) SDS-PAGE. Protein bands were visualized by staining with Coomassie Brilliant Blue R-250. Using a protein prestained ladder from Fermentas as standard.

B) Electrospray ionization mass spectrometry of proteins. Sample preparation: each protein (1 mg of lyophilized powder) was resuspended in water (1 mL) and formic acid (5 μL) was added. Samples (10 μL) were analyzed by HPLC-ESI-MS using an Acquity UPLC^{TM} BEH300 C18 column (1.7 μm, 2.1 x 100 mm), and an ESI-TOF mass spectrometer (LCT PremierWaters, Milford, MA, USA) equipped with a 4GHz time-to-digital converter (TDC) with a dual ESI source (LockSpray). The second sprayer provided the lock mass calibration with leucine enkephalin (m/z 556.2771). The ESI-TOF was operated in the W-optics mode, thus providing a mass resolution of at least 10 000 full-width at half maximum (FWHM). The acquisition time per spectrum was set to 0.2 s, and the mass range was from 500 to 1800 Da. Data were acquired using a cone voltage of 50 V, capillary voltage of 3000 V, desolvation temperature of 350 ºC, and source temperature of 100 ºC. The desolvation gas flow was set at 400 L h^{-1} and the cone gas flow was set at 30 L h^{-1}. The solvent system used for the elution was: solvent (A): aqueous

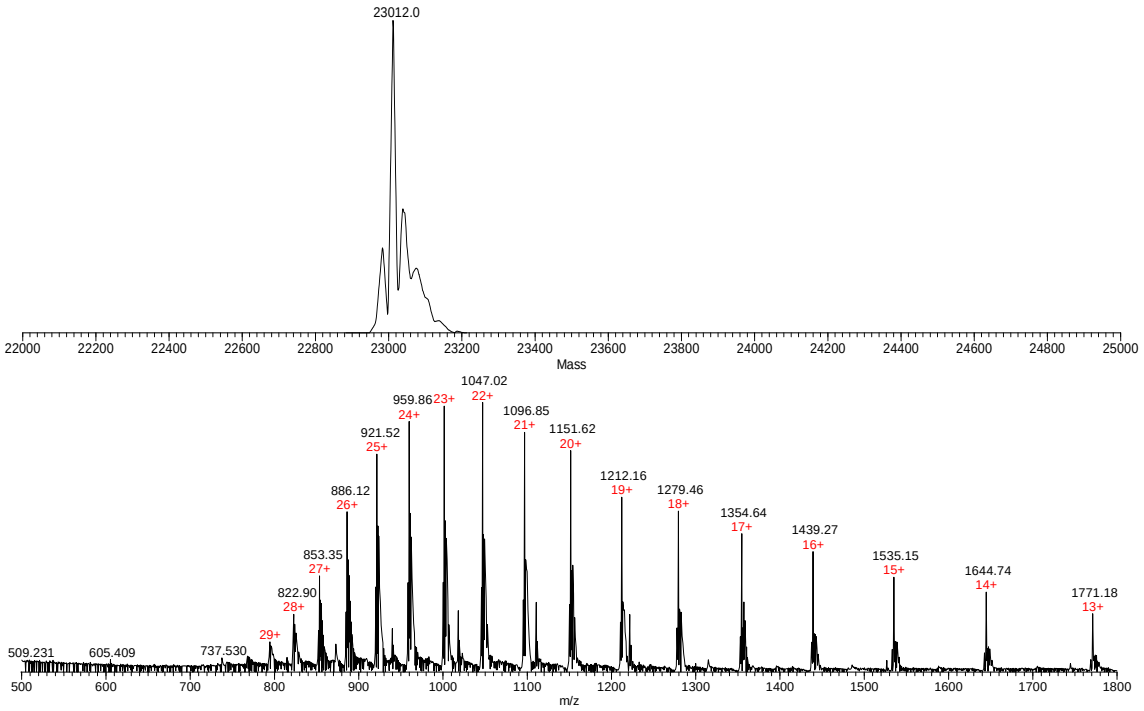
formic acid 0.1 % (v/v) and solvent (B): formic acid 0.1 % (v/v) in acetonitrile, gradient elution 0 % B for 5 min, from 5 to 70 % B in 12 min, from 70 to 100 % B in 1 min, flow rate 0.3 mL min^{-1}. MassLynx 4.1 (Waters. Milford, MA, USA) was used for data acquisition and processing. Magtran software[7] kindly provided by Dr. Zhongqi Zhang (Amgen, Inc.; Thousand Oaks, CA) was used for molecular weight deconvolution from ESI-MS spectra of proteins (Figure S1 and Table S2).

Table S2. Mass spectrometry of the FSA mutants.

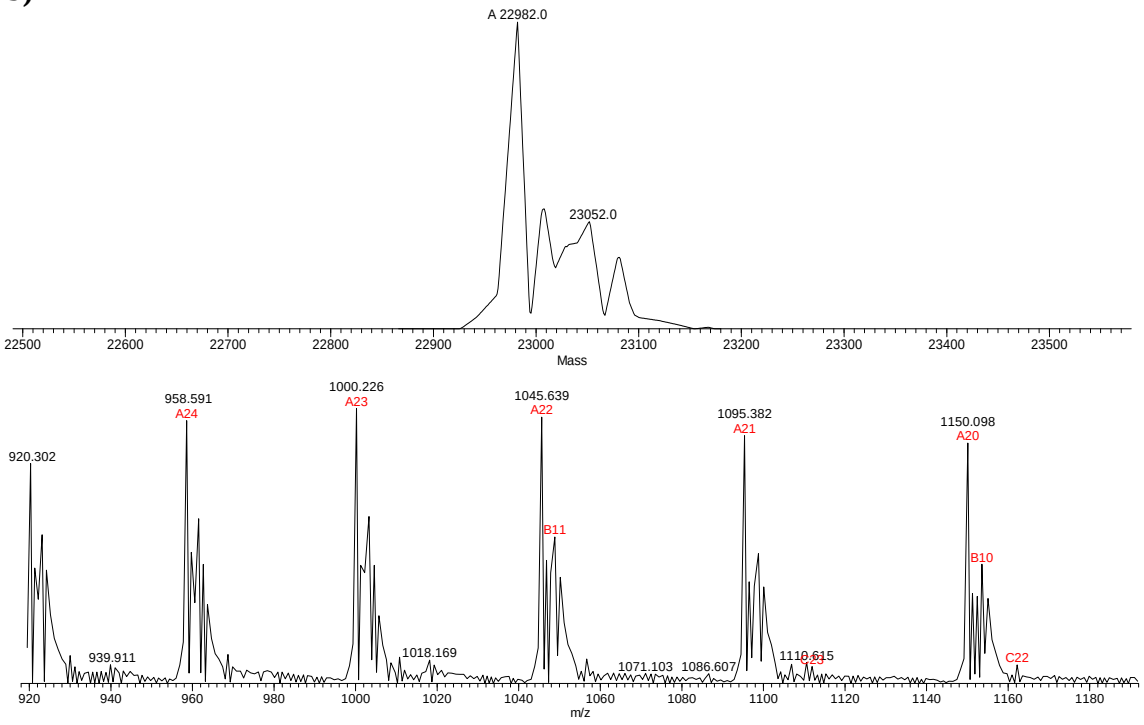
FSA mutants	**Calculated**	**Observed**	Error abs (Da)
FSA A129S	23012.82	23012.0	0.82
FSA A165G	22982.80	22982.0	0.80
FSA A129S A165G	22998.80	22999.0	0.20
FSA R134A	22911.71	22911.0	0.71

Figure S2. MS ESI/TOF spectra and deconvolution spectra for the FSA mutants. a) FSA A129S, b) FSA A165G, c) FSA A129S A165G, d) FSA R134A.

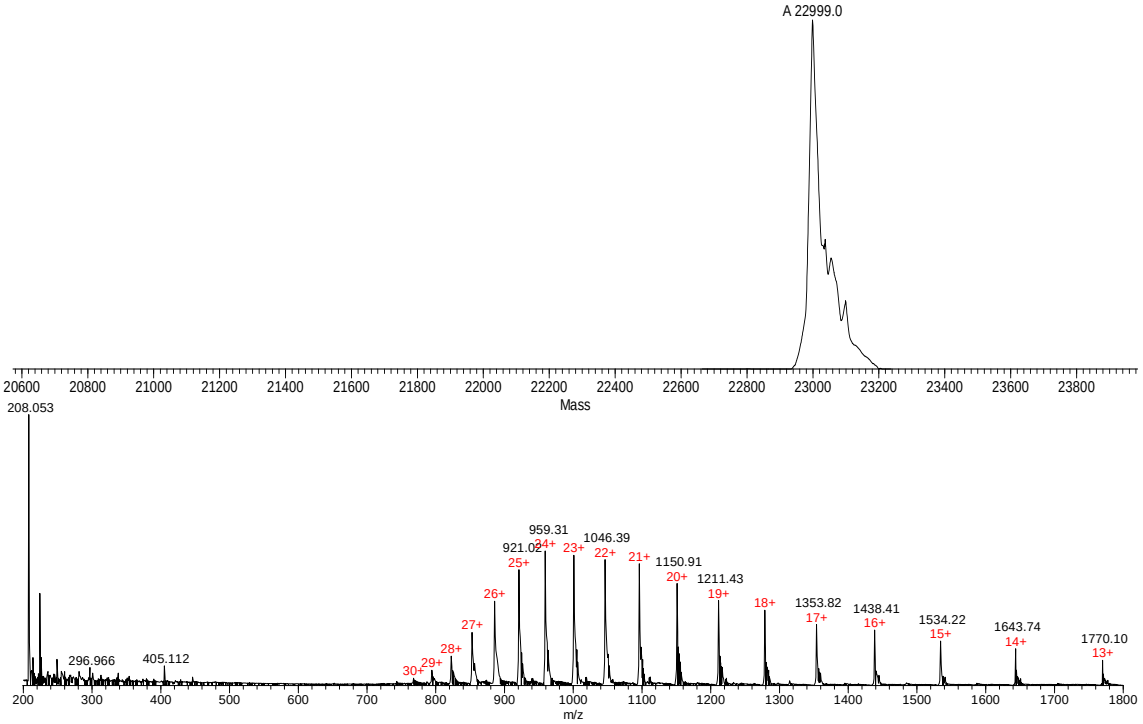
a)



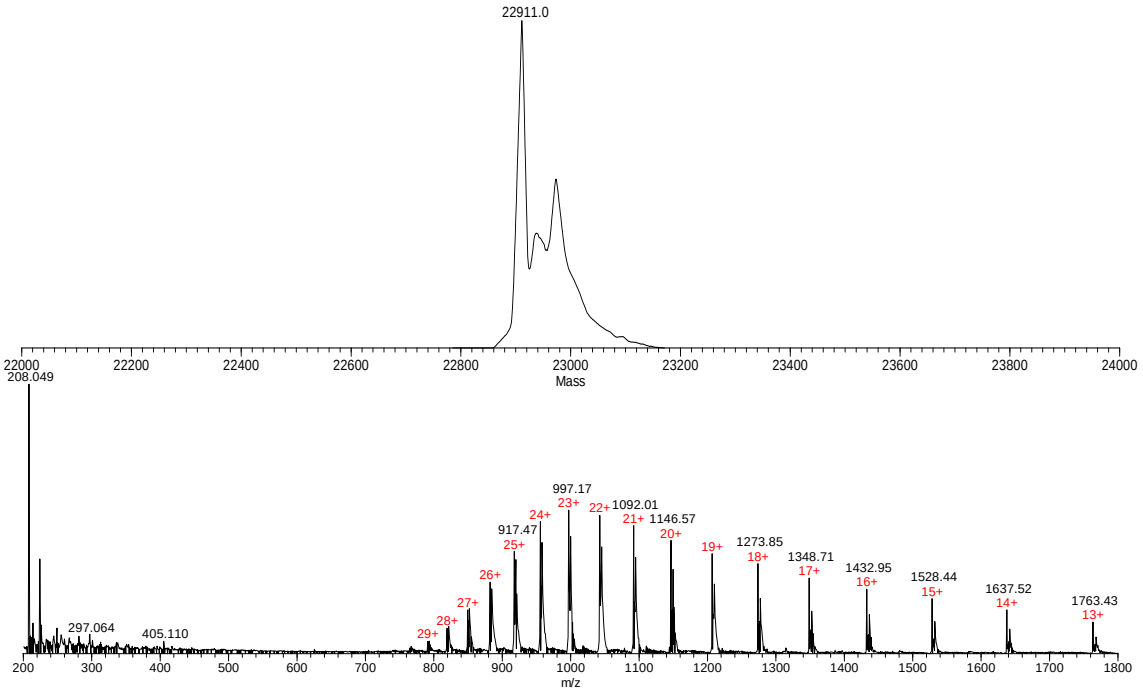
b)



c)


A 22999.0
Mass
208.053
296.966
405.112
959.31
921.02
1046.39
1150.91
1211.43
1353.82
1438.41
1534.22
1643.74
1770.10
m/z

d)


22911.0
Mass
208.049
297.064
405.110
997.17
917.47
1092.01
1146.57
1273.85
1348.71
1432.95
1528.44
1637.52
1763.43
m/z

C) Specific activity of the mutants on the aldol addition reaction of dihydroxyacetone phosphate to D,L-glyceraldehyde-3-phosphate.

The specific activity of the FSA wild-type and mutants was determined using the assay described above consisting of the synthesis of D-fructose-6-phosphate from the aldol addition of DHA and D,L-glyceraldehyde-3-phosphate (G3P). As expected, the lowest activity was found for the FSA R134A since the acceptor G3P loss the ionic interaction site with the R134 residue (Table S3). Furthermore, the A129S mutant was the most active, in good agreement with a previous work,[8] and the double A129S/A165G had higher activity than that of A165G due to the putative additional stabilization of the DHA with the S129 residue (Table S3).

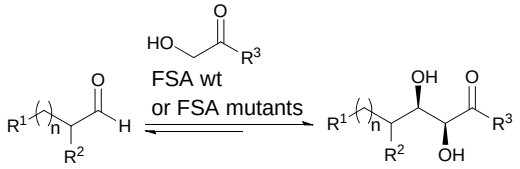
Table S3 Specific aldolase activity[a] of the FSA wild-type and the mutants used in this work.

Aldolase	Specific U mg^{-1}protein (Bradford)
FSA	1.5±0.2
FSA R134A	0.3±0.1
FSA A129S	9.4±1.3
FSA A165G	0.9±0.1
FSA A129S/A165G	2.5±0.9

[a]One Unit (U) catalyzed the formation of 1 μmol of D-fructose-6-phosphate (D-F6P) per minute from DHA (300 mM) and D,L-glyceraldehyde-3-phosphate (D,L-GAP) (2.8 mM) at 30 ºC in glycylglycine buffer (1 mL, 50 mM, pH 8.5) containing DTT (1 mM).[5] Values are the average of at least three independent measurements.

D) Enzymatic aldol reactions

Enzymatic aldol reactions with DHA or HA as donors (Scheme S1): Analytical scale reactions. A preliminary set of reactions was conducted to test the obtained mutant proteins with selected *N*-Cbz-amino aldehydes (Table S4). Reactions (200 μL total volume) were performed using 1 mL test tubes stirred with a vortex mixer (VIBRAX VXR basic, Ika) at 800 rpm and 25 ºC. Each of the aldehydes (0.016 mmol, 80 mM) was dissolved in 40 μL of dimethylformamide. Then, to this solutions, freshly prepared DHA or HA solution (0.02 mmol) in 160 μL of 50 mM glycylglycine buffer pH 8, containing DTT (1 mM), at 25 ºC was added. The reactions were started by adding the enzyme (0.5 mg protein, for the units see Table S3). At 0, 4 and 24 hours samples (25 μL) were withdrawn and dissolved in methanol (475 μL), and then subsequently analyzed by HPLC.



n = 0	R_1	R_2
(*S*)-**1**	-NH-Cbz	(*S*)CH_3
(*R*)-**1**	-NH-Cbz	(*R*)CH_3
2	-NH-Cbz	H
3	-O-Bn	H
n = 1		
4	-NH-Cbz	H

n = 0	R_1	R_2	R_3
(5*S*)-**5**	-NH-Cbz	(*S*)CH_3	CH_2OH
(5*S*)-**6**	-NH-Cbz	(*R*)CH_3	CH_3
(5*R*)-**5**	-NH-Cbz	(*S*)CH_3	CH_2OH
(5*R*)-**6**	-NH-Cbz	(*R*)CH_3	CH_3
7	-NH-Cbz	H	CH_2OH
8	-NH-Cbz	H	CH_3
9	-O-Bn	H	CH_2OH
10	-O-Bn	H	CH_3
n = 1			
11	-NH-Cbz	H	CH_2OH
12	-NH-Cbz	H	CH_3

Scheme S1. FSA as catalyst for the aldol addition of DHA and HA to aldehydes **1**-**4**.

Table S4. Initial screening of the FSA wild-type and mutants with the selected aldehydes.

		FSA wild-type	FSA A129S	FSA A165G	FSA R134A	FSA A129/A165G
Acceptor aldehyde	Donor	Aldol adduct formed (%)[a]				
(*S*)-1	DHA	nd	nd	37	nd	75
	HA	35	21	94	45	>98[b]
(*R*)-1	DHA	nd	nd	65	nd	90
	HA	nd	nd	90	nd	92
2	DHA	20	43	95	nd	>98[b]
	HA	92	97	95	93	>98[b]
3	DHA	8	40	63	10	>98[b]
	HA	90	90	92	87	>98[b]
4	DHA	68	87	85	60	94
	HA	86	90	90	75	96

[a]Percentage of aldol adduct formed after 24 h measured by HPLC using an external standard method. Conditions: aldehyde (0.016 mmol, 80 mM), DHA (0.02 mmol, 80 mM) or HA (0.02 mmol, 100 mM) and the enzyme (0.5 mg) in 200 μL total volume. [b]no aldehyde left was detected by HPLC; limit of detection was between 0.2 and 0.5 nmol. nd = aldol adduct not detected by HPLC.

The most productive enzymes were chosen for a second set of reactions to calculate the initial reaction rates (v_o, mmol min^{-1} mg^{-1} protein) and maximum conversions. Reactions (300 μL total volume) were performed as described above using 1.5 mL Eppendorf type test tubes: the aldehyde acceptor (0.024 mmol, 80 mM) was dissolved in dimethylformamide (60 μL). Then, DHA or HA (0.030 mmol) in 210 μL of 50 mM glycylglycine buffer pH 8, containing DTT (1 mM) was added and equilibrated at 25 ºC. The reaction was initiated by the addition of the corresponding FSA mutant and mixing. The amount of enzyme varied from 0.05 to 1 mg of protein (for Units see Table 1 in the paper) to ensure a linear dependence of the product formation versus time at the beginning of the reaction (reaction conversions <15%). Reactions were placed in a vibrax mixer (800 rpm) at 25 ºC. At different reaction time (FSA wild-type and FSA A129S: conversion 0 h

and 24 h; v_o between 0 and 120 min; FSA A165G and FSA A129S/A165G conversion 0 and 24 h;v_o between 0 and 30 min), samples (25 μL) were withdrawn, mixed with methanol (975 μL) centrifuged to removed the precipitated protein, and analyzed directly by HPLC. Concentration of products was calculated from the peak areas using an external standard method. Conversions were calculated with respect to the initial concentration of aldehyde acceptor, which were the limiting reactants. Linear correlations were found for conversion <15 %. HPLC conditions and sample concentration were adjusted to assure minimum error with the external standard methodology.

The stereochemical outcome of the reactions, particularly with the FSA A129/A165G double mutant, was assessed by following our previously reported methodologies.[1] Analytical and preparative scale reactions were performed for the selected aldehydes, and the corresponding aldol adducts converted into iminocyclitols. The NMR spectra of the crude were compared to those previously published with FSA wild-type and D-fructose-1,6-bisphophate aldolase (see below).[1,3,9,10]

E) Steady-state kinetic parameters for DHA and *N*-Cbz-3-aminopropanal (4). To determine the K_m^{app} and V_{max}^{app} values for DHA and *N*-Cbz-3-aminopropanal for the aldol reaction of DHA to **4** catalyzed by FSA wild-type, FSA A129S, FSA A165G and the double mutant FSA A129S/A165G, analytical scale reactions (total volume was 300 μL) were conducted in eppendorf type tubes (1.5 mL), stirred with a vortex mixer (VIBRAX VXR basic, Ika) at 800 rpm and 25 ºC.

Steady-state kinetic parameters for DHA. Aldehyde **4** (30 μmol) was dissolved in dimethylformamide (60 μL) and mixed with a solution (190 μL) containing different amounts of DHA (6, 12, 18 and 45, μmol for FSA wild-type, FSA A129S and FSA A129S/A165G and 6, 9, 12, 18 , 30 and 45, μmol for FSA A165G) in glycylglycine buffer (50 mM pH 8) containing DTT (1 mM). To this mixture, the aldolase, FSA wild-type, FSA A129S, FSA A165G or FSA A129S A165G (50 μL, containing a range of protein between 0.1 to 0.85 mg depending on the reaction) was added to start the reaction. For each DHA concentration, samples (25 μL) were withdrawn at different times (in a range between 0-30 min depending on the reaction), diluted with methanol (475 μL), centrifuged and analyzed by HPLC as described for the determination of initial velocities. The experimental and fitted data of the aldolase activity vs DHA concentration, was adjusted

by computer-based nonlinear regression of untransformed data to the kinetic Michaelis-Menten equation using GraphPad Prism software (see Figure S2)

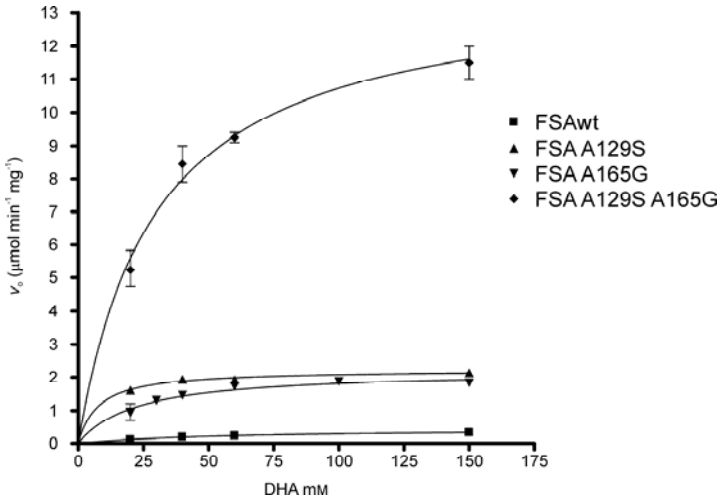


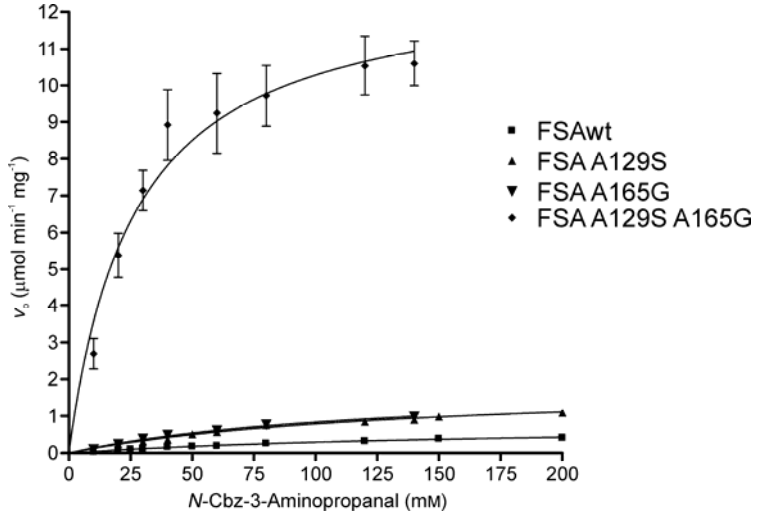
Figure S3. Activity (μmol of aldol adduct formed min^{-1} mg^{-1} of protein) of FSA wild-type (■), FSA A129S (▲), FSA A165G (▼) and FSA A129S/A165G (◆) at different DHA concentrations for the aldol addition of DHA to **4**. Experimental data was adjusted (solid lines) by a computer non-linear regression method to the Michaelis-Menten equation. Standard deviations were calculated from experiments performed by triplicate.

Steady-state kinetic parameters for N-Cbz-3-aminopropanal. DHA (90 μmol) was dissolved in 190 μL of glycylglycine buffer (50 mM pH 8) containing DTT (1 mM), and mixed with a solution of dimethylformamide (60 μL) containing different amounts of *N*-Cbz-3-aminopropanal (3, 6, 7.5, 9, 12, 18, 24, 36, 45 and 60 μmol for FSA wild-type; 3, 6, 9, 12, 15, 18, 24, 36, 45 and 60 μmol for FSA A129S; 3, 6, 9, 12, 18, 24 and 42 μmol for FSA A165G and 3, 6, 9, 12, 18, 24, 36 and 42 for FSA A129S/A165G). To this mixture, the aldolase, FSA wild-type, FSA A129S, FSA A165G or FSA A129S A165G, (50 μL, containing a range of protein between 0.1 to 0.85 mg depending on the reaction) was added to start the reaction (total final volume, 300 μL). For each reaction, samples (25 μL) were withdrawn at different times (in a range between 0-30 min depending on the reaction), diluted with methanol (475 μL), centrifuged and analyzed by HPLC as described above.

The experimental and fitted data of the aldolase activity vs *N*-Cbz-3-aminopropanal concentration, was adjusted by computer-based nonlinear regression of untransformed

data to the kinetic Michaelis-Menten equation using GraphPad Prism software (see Figure S3)

a)



b)

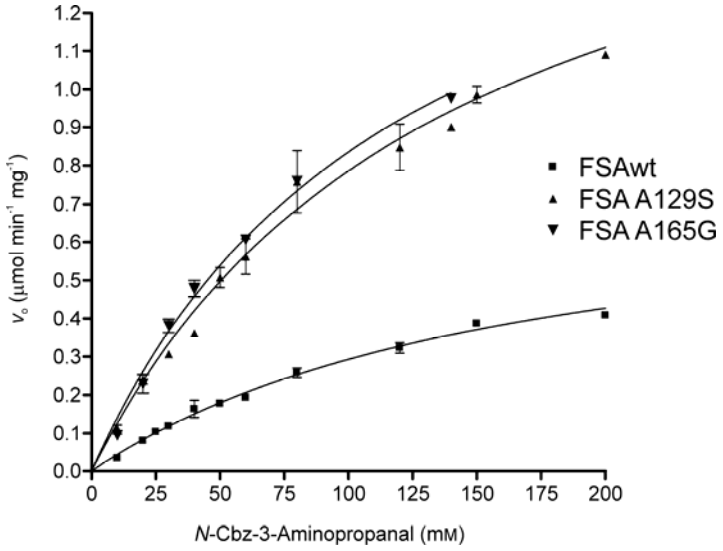
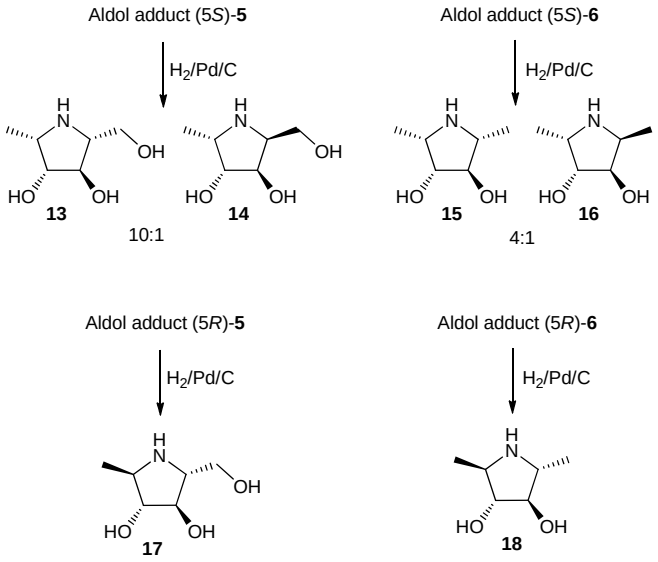


Figure S4. Activity (µmol of aldol adduct formed min^{-1} mg^{-1} of protein) of FSA wild-type (■), FSA A129S (▲), FSA A165G (▼) and FSA A129S/A165G (◆) at different *N*-Cbz-3-aminopropanal (**4**) concentrations for the aldol addition of DHA to **4**. Curve a) for FSA wild-type type and mutants, curve b) only for FSA wild-type, FSA A129S and FSAA165G. Experimental data was adjusted (solid lines) by a computer non-linear regression method to the Michaelis-Menten equation. Standard deviations were calculated from experiments performed by triplicate.

F) Scale up of the aldol addition of DHA and HA to aldehydes catalyzed by FSA A129S/A165G. General procedure. Reactions were conducted in 50 mL screw capped conical-bottom polypropylene test tubes with 12.5 mL total reaction volume using the same reactant and enzyme concentration and procedure used in the corresponding analytical reactions: aldehydes (1 mmol) and DHA or HA (1.25 mmol) in 50 mM $NaHCO_3$-Na_2CO_3 buffer pH 7.5. When the product concentration was constant by HPLC (~24h, conversions similar to those described for analytic reactions), a simple work up/purification on Amberlite™ XAD™16 column was carried out. The fractions were analyzed by HPLC. Pure fractions were pooled and lyophilized obtaining the corresponding aldol adducts (5*S*)-**5** (248 mg, 0.83 mmol, 83% isolated yield), (5*S*)-**6** (257 mg, 0.91 mmol, 91% isolated yield), (5*R*)-**5** (214 mg, 0.72 mmol, 72% isolated yield), (5*R*)-**6** (220 mg, 0.78 mmol, 78% isolated yield); aldol adduct **7** (**2** (14 mmol) and DHA (18 mmol), 2.3 g, 82 % isolated yield). The aldol adducts **8**, **9**, **10**, **11** and **12** were prepared at analytical scale (2 mL, 10-20 mg of pure material). The aldol adducts (5*S*)-**5**, (5*S*)-**6,** (5*R*)-**5**, (5*R*)-**6**, **7**, **8**, **11** and **12** were converted into the corresponding iminocyclitols. To this end, products were dissolved in H_2O/MeOH 1:1 and treated with H_2 (50 psi) in the presence of Pd/C (100 mg) at room temperature during 24 h. After removal of the catalyst by filtration through 0.45 μm nylon membrane filter, the filtrates were adjusted to pH 6.4 with formic acid and lyophilized obtaining an oil-like material, which was not further purified. This material was analyzed by NMR and the iminocyclitol structures were unequivocally assigned. The ^{1}H, ^{13}C NMR spectra of the iminocyclitols from aldol adducts (5*R*)-**5**, (5*S*)-**5**, **7**, **8**, and **12** matched those previously reported using D-fructose-1,6-bisphosphate aldolase from rabbit muscle (RAMA) and FSA wild-type.[1,3,10] ^{1}H and ^{13}C NMR of iminosugar from aldol adduct **12** matched that reported by Sugiyama et al.[10] ^{1}H and ^{13}CNMR of aldol adducts **9, 10** and **11** were indistinguishable from those reported previously by us.[1] Specific rotation of the iminocyclitols from the aldol adducts **7**, **8** and **11** were coincident to those reported and obtained using similar experimental procedure but using FSA wild-type.[1]

The physical and spectral data of the iminocyclitols prepared using the FSA A129S/A165G that could not be previously prepared by FSA wild-type are the following (Scheme S2):



Scheme S2. Formation of iminocyclitols **13**-**18** from aldol adducts (5*S*)-**5**, (5*R*)-**5,** (5*S*)-**6** and (5*R*)-**6** by means of deprotection of the Cbz group and reductive amination with H_2 in the presence of Pd/C.

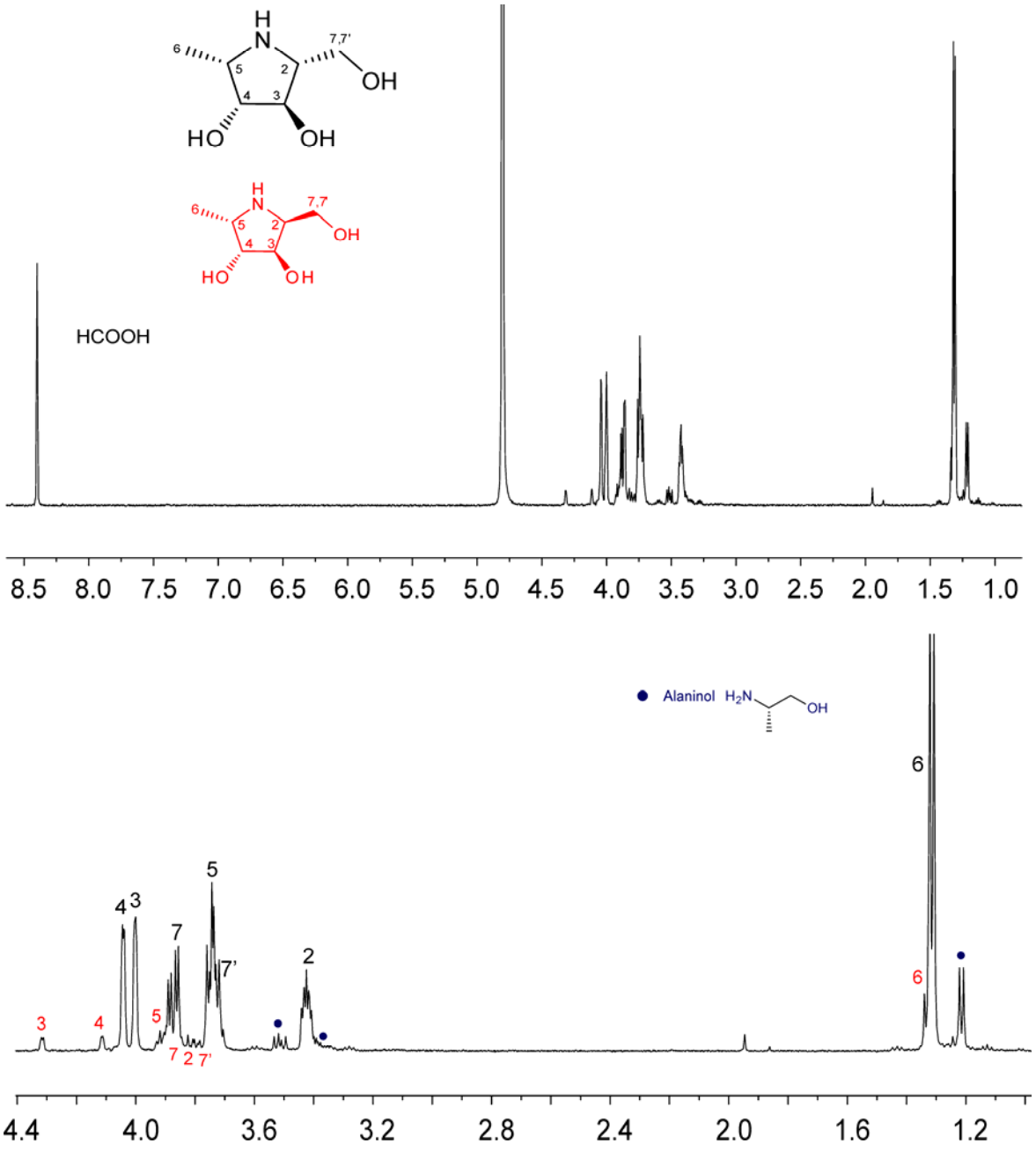
Specific rotation of the iminocyclitols **13**+**14** (as a mixture 10:1, arising from the reductive amination step) and **17** from the aldol adducts (5*S*)-**5** and (5*R*)-**5**, respectively as formiate salts: **13**:**14** 10:1; $[\alpha]_D^{22}$ = + 32.8 (*c*=0.7 in MeOH) (lit.[3] $[\alpha]_D^{22}$ = + 16.0 (*c*=0.4 in MeOH); lit.[11] $[\alpha]_D^{22}$ = – 43.30 (C=0.8 in H_2O); lit.[12] $[\alpha]_D^{22}$ = + 22.7 (*c*=1.2 in MeOH)) and **17;** $[\alpha]_D^{22}$ = + 26.7 (*c*=0.6 in MeOH) (lit.[3] $[\alpha]_D^{22}$ = + 18.7 (*c*=1 in MeOH); lit.[13] $[\alpha]_D$ = + 26.2 (c=1.1, in MeOH)] for the natural compound; lit.[12] $[\alpha]_D^{22}$ = + 39.1 (*c*=0.8 in MeOH)). ^{1}H, ^{13}C NMR spectra matched those already published (see below for a copy of the spectra).[3,12,13]

^{1}H, ^{13}C NMR spectra (see below for a copy of the spectra) and the specific rotation of the iminocyclitols **15**+**16** (as a mixture 4:1 arising from the reductive amination step) and **18** from aldol adduct (5*S*)-**6** and (5*R*)-**6**, respectively. **15**+**16** (as a mixture 4:1) $[\alpha]_D^{22}$ = + 18.3 (*c*=0.93 in H_2O). (2*R*,3*R*,4*R*,5*S*)-2,5-Dimethylpyrrolidine-3,4-diol (**15**) (79%): ^{1}H NMR (500 MHz, D_2O) δ = 4.03 (m, 1H), 3.94 (dd, *J* = 3.1, 1.7 Hz, 1H), 3.77 (m, 1H), 3.46 (m, 1H), 1.40 (d, *J* = 7.2 Hz, 3H), 1.30 (d, *J* = 6.8 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ = 83.0, 79.1, 64.2, 60.8, 18.8, 13.0. (2*S*,3*R*,4*R*,5*S*)-2,5-Dimethylpyrrolidine-3,4-diol (**16**) (21%) ^{1}H NMR (500 MHz, D_2O) δ = 4.11 (d, *J* = 3.0 Hz, 1H), 3.86 (m, 1H), 1.28 (d, *J* = 7.0 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ = 78.6, 59.9, 13.0.

(2*R*,3*R*,4*R*,5*R*)-2,5-Dimethylpyrrolidine-3,4-diol (**18**). $[\alpha]_D^{22}$ = + 26.7 (*c*=0.3 in H_2O). ^1H NMR (500 MHz, D_2O) δ = 3.74 (m, 1H), 3.39 (m, 1H), 1.31 (d, *J* = 7.0 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ = 82.48 (s), 59.53 (s), 17.92 (s).

Figure S5. Observed NMR spectra of the iminocyclitols **13** (major product, i.e. 2,5-imino-1,2,5-trideoxy-D-glucitol) and **14** obtained from the aldol adduct (5*S*)-**5** mediated by FSA A129S/A165G double mutant. a) ^{1}HNMR, b) ^{13}C.

a)



b)

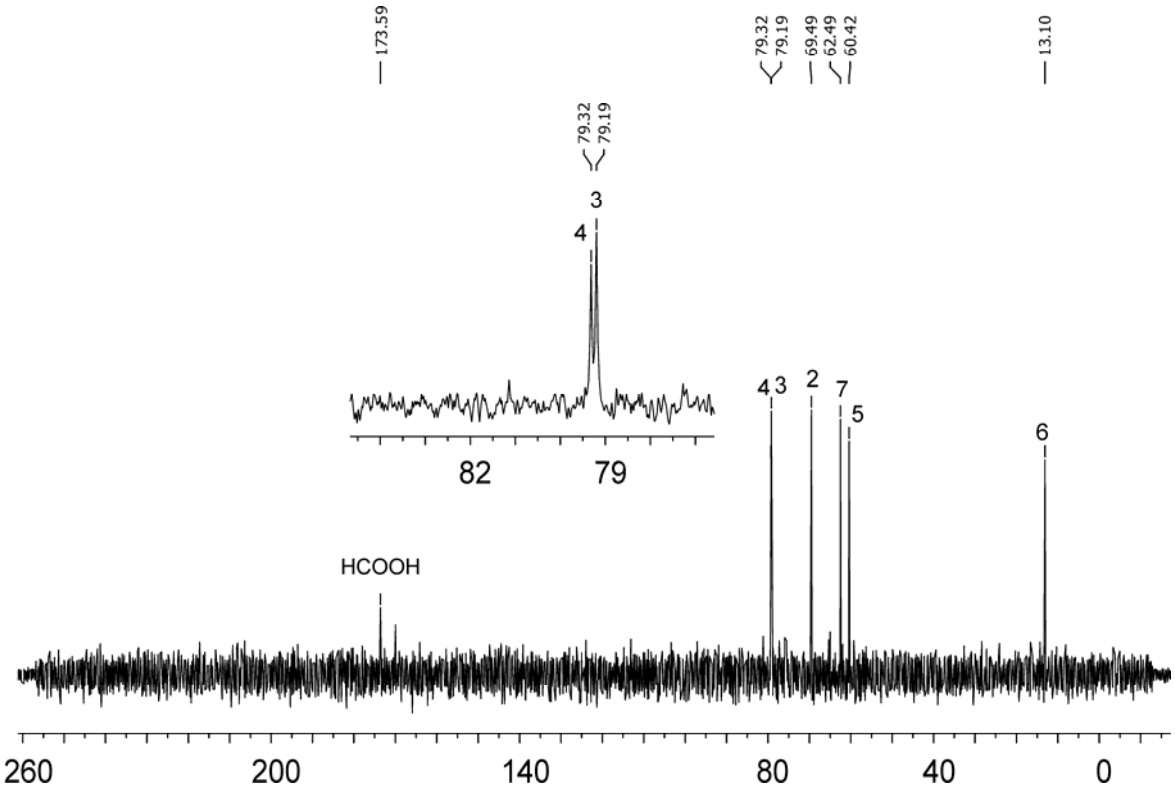
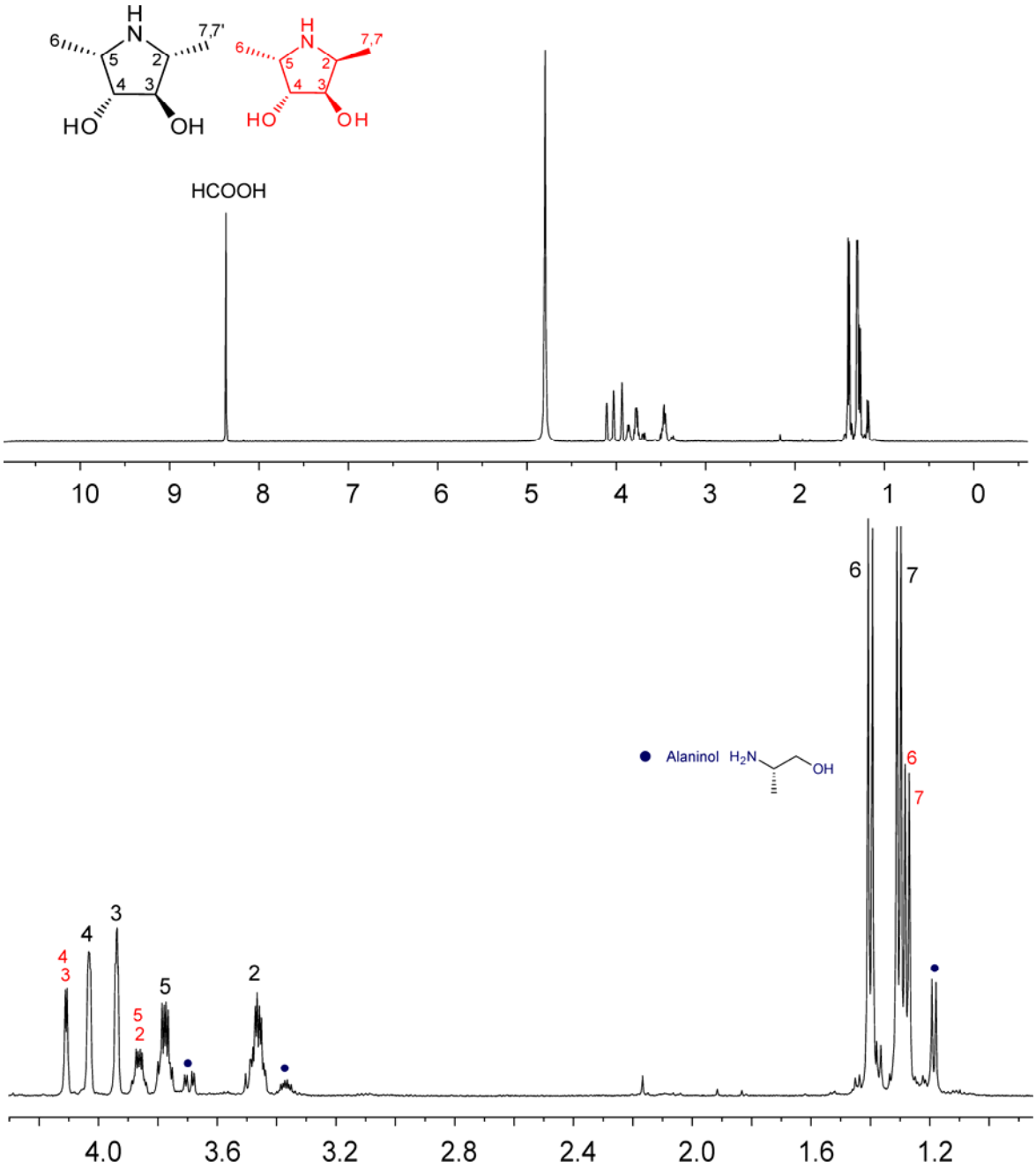
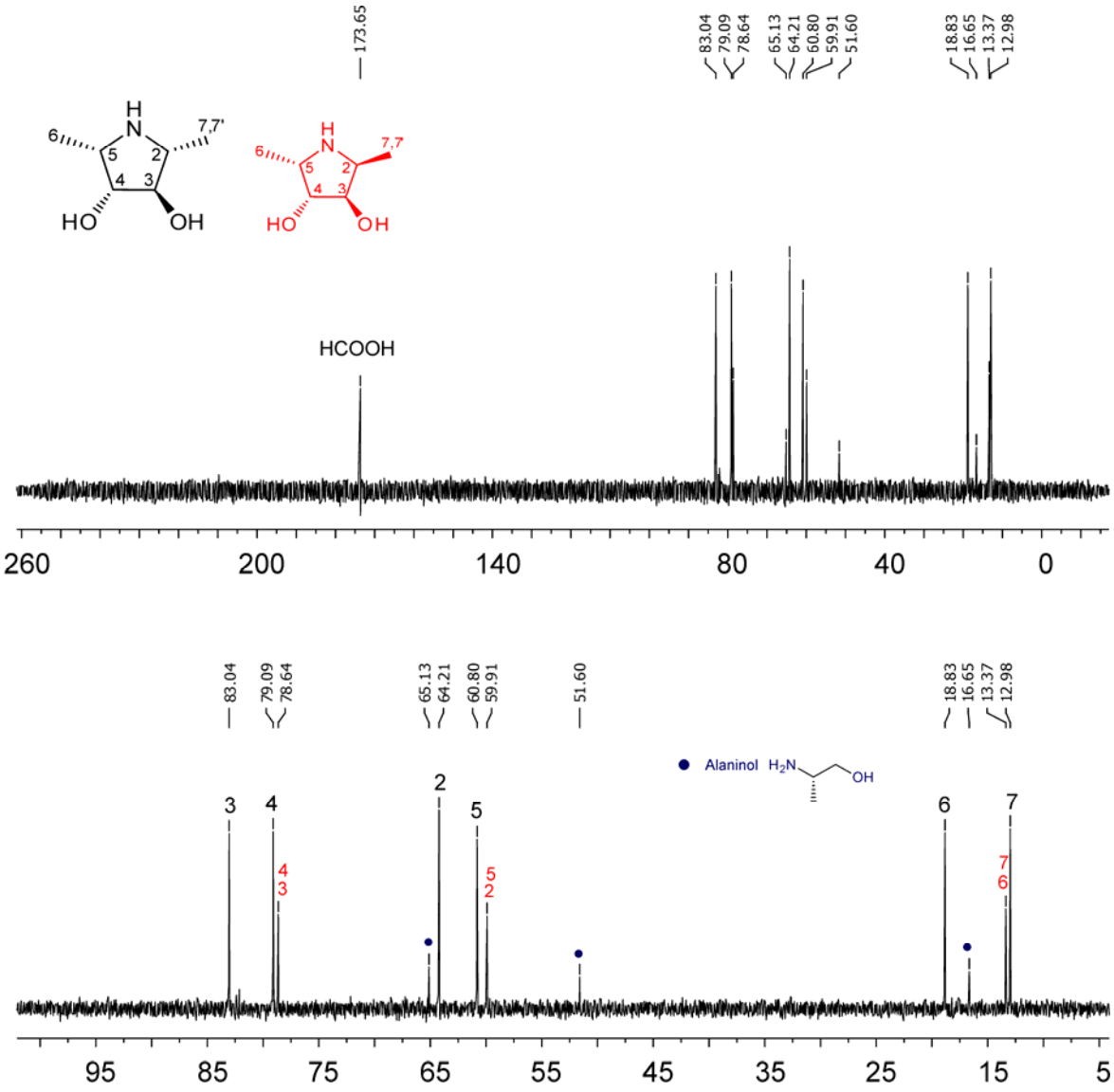

173.59
79.32
79.19
69.49
62.49
60.42
13.10
79.32
79.19
3
4
82
79
4 3
2
7
5
6
HCOOH
260
200
140
80
40
0

Figure S6. Observed NMR spectra of the iminocyclitols **15** (major) and **16** (minor) obtained from the aldol adduct (5*S*)-**6** mediated by FSA A129S/A165G double mutant. a) ^{1}HNMR, b) ^{13}C, c) 2D 1H-1H COSY and 2D multiplicity-edited HSQC and d) selective 1D NOESY spectra.

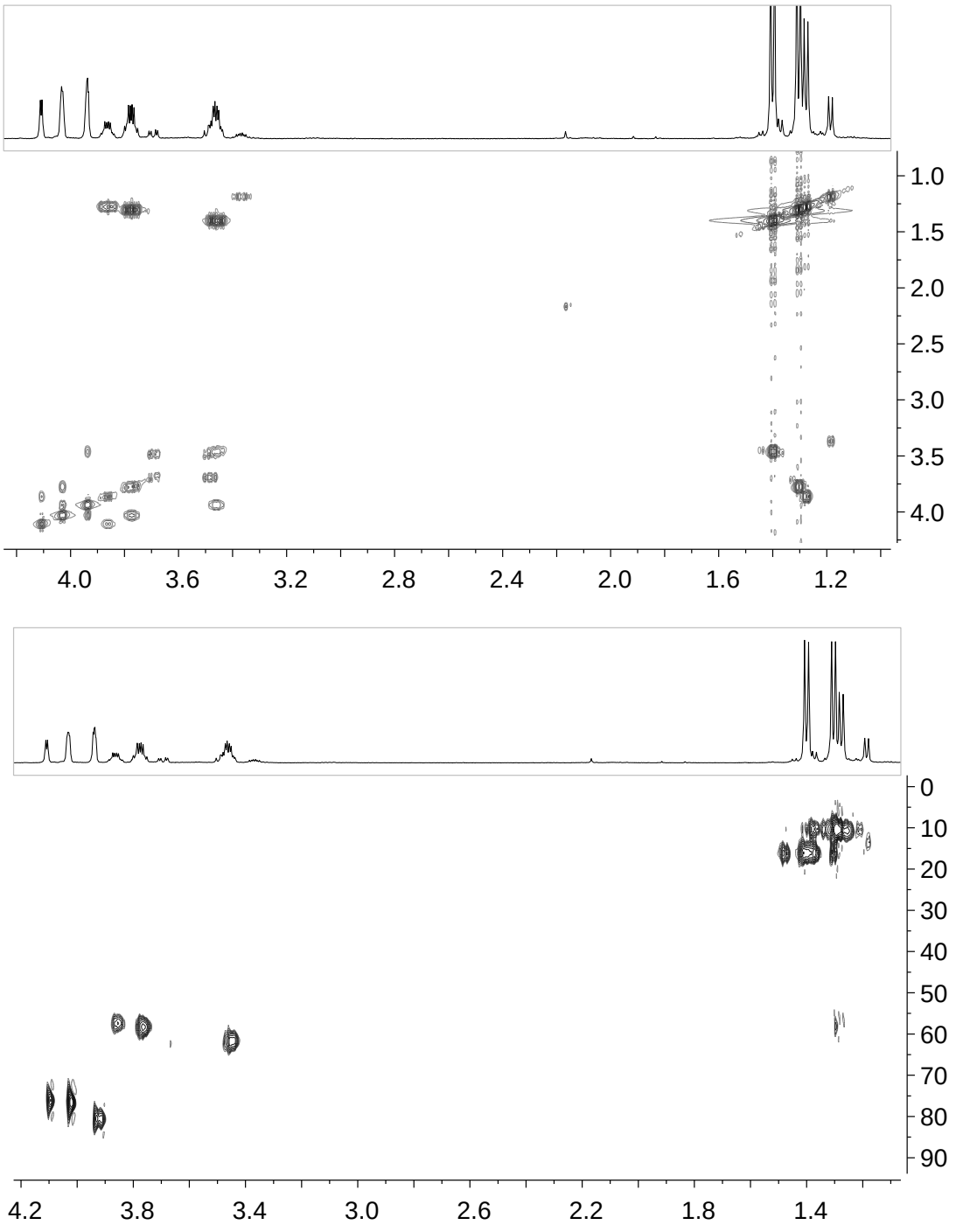
a)



b)


173.65
83.04
79.09
78.64
65.13
64.21
60.80
59.91
51.60
18.83
16.65
13.37
12.98
HCOOH
260
200
140
80
40
0
Alaninol
95
85
75
65
55
45
35
25
15
5

c)


1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0
10
20
30
40
50
60
70
80
90
4.2
3.8
3.4
3.0
2.6
2.2
1.8
1.4

d)

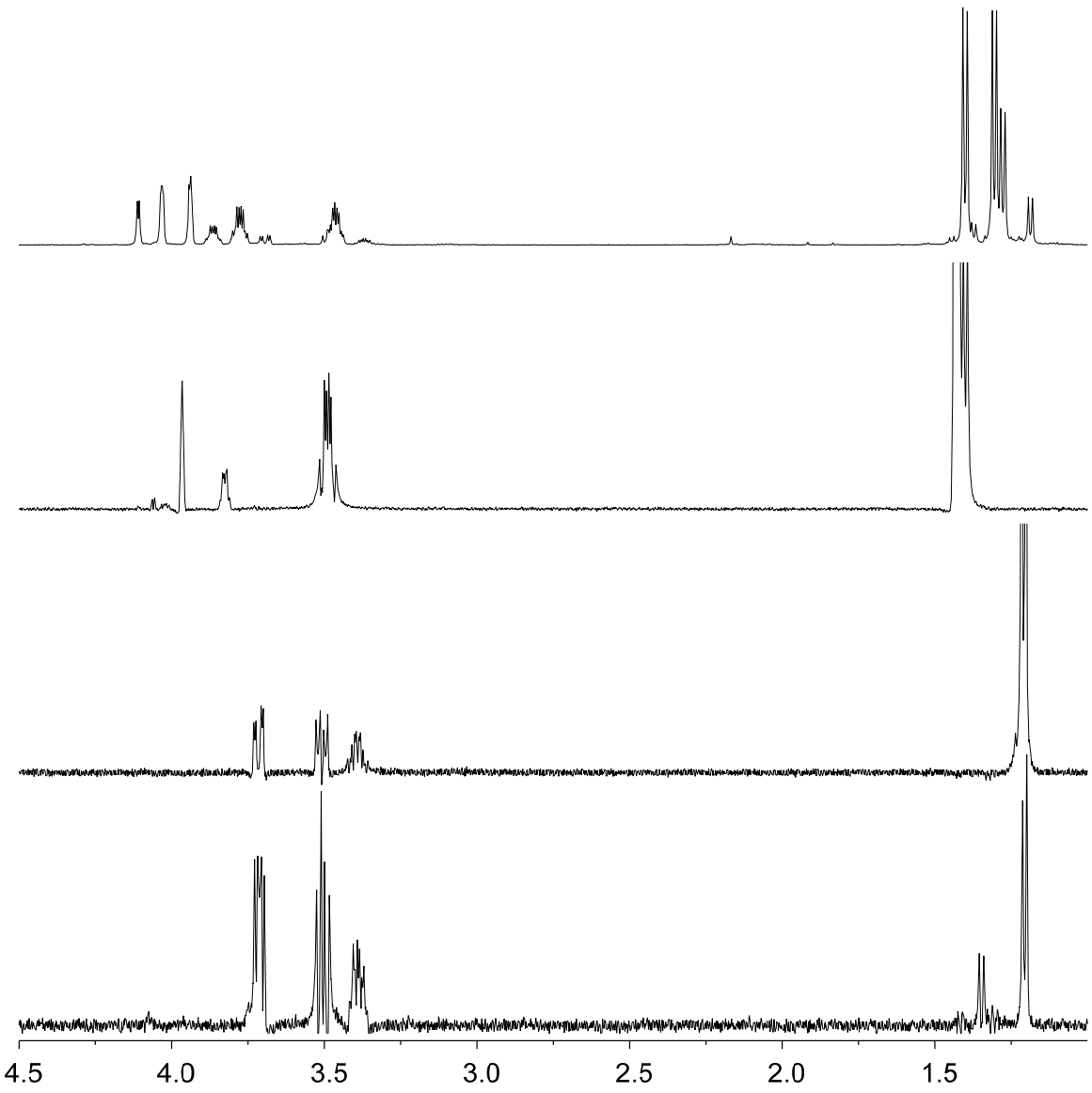
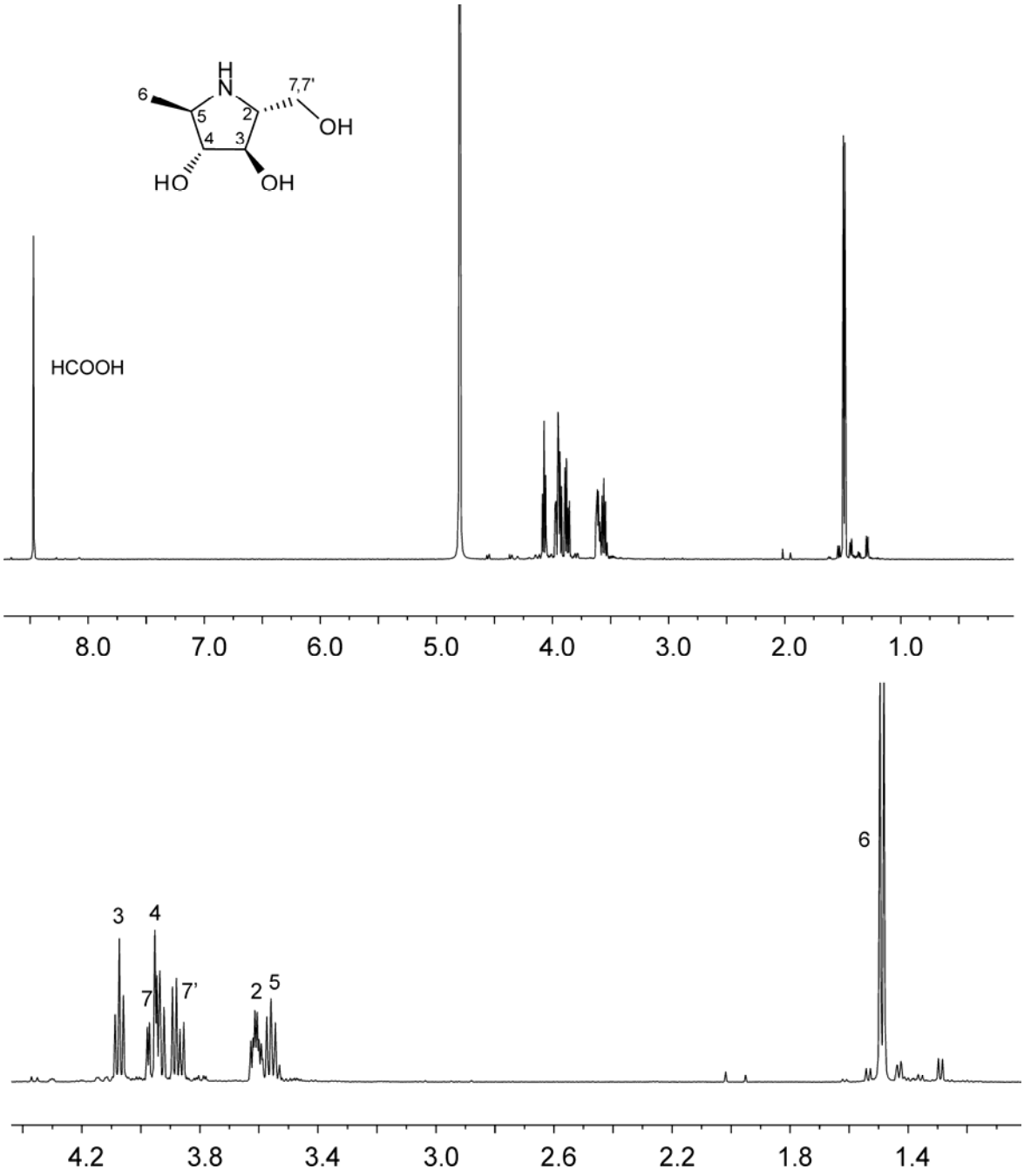


Figure S7. NMR spectra of 2,5-imino-1,2,5-trideoxy-D-mannitol (6-deoxy-DMDP) **17** obtained from the aldol adduct (5*R*)-**5** mediated by the FSA A129S/A165G double mutant. a) ^{1}HNMR and b) ^{13}CNMR.

a)



b)

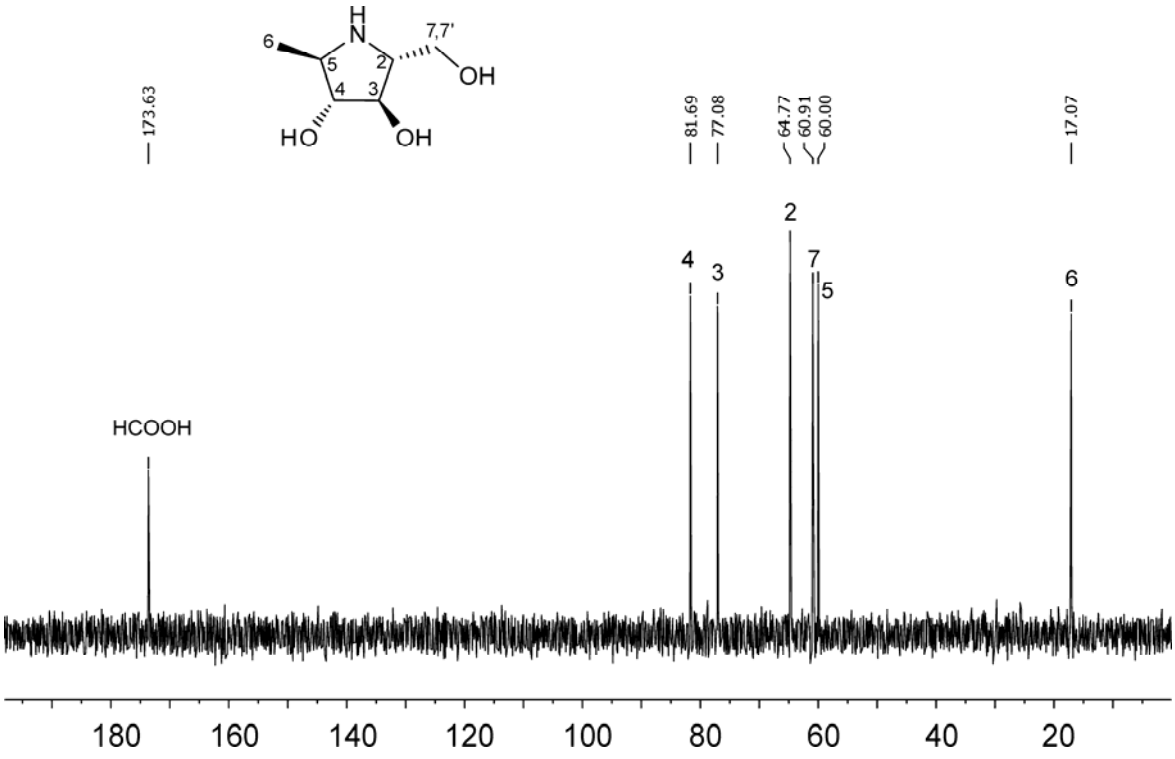

H
N
6
5
2
7,7'
OH
4
3
HO
OH
173.63
81.69
77.08
64.77
60.91
60.00
17.07
HCOOH
2
4
3
7
5
6
180
160
140
120
100
80
60
40
20

Figure S8. Observed ^{1}H and ^{13}CNMR spectra of the iminocyclitol **18** obtained from the aldol adduct (5*R*)-**6** mediated by the FSA A129S/A165G double mutant.

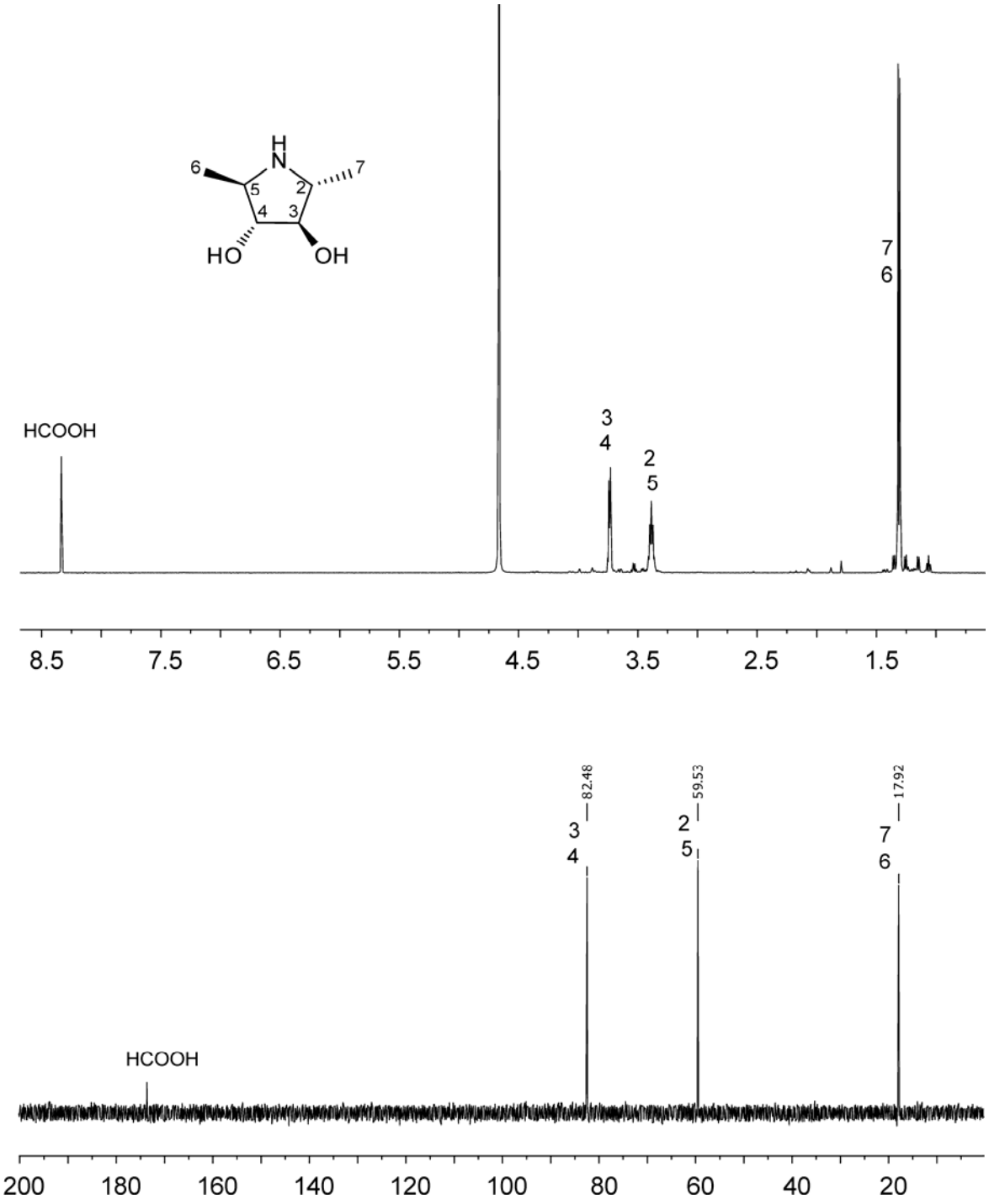
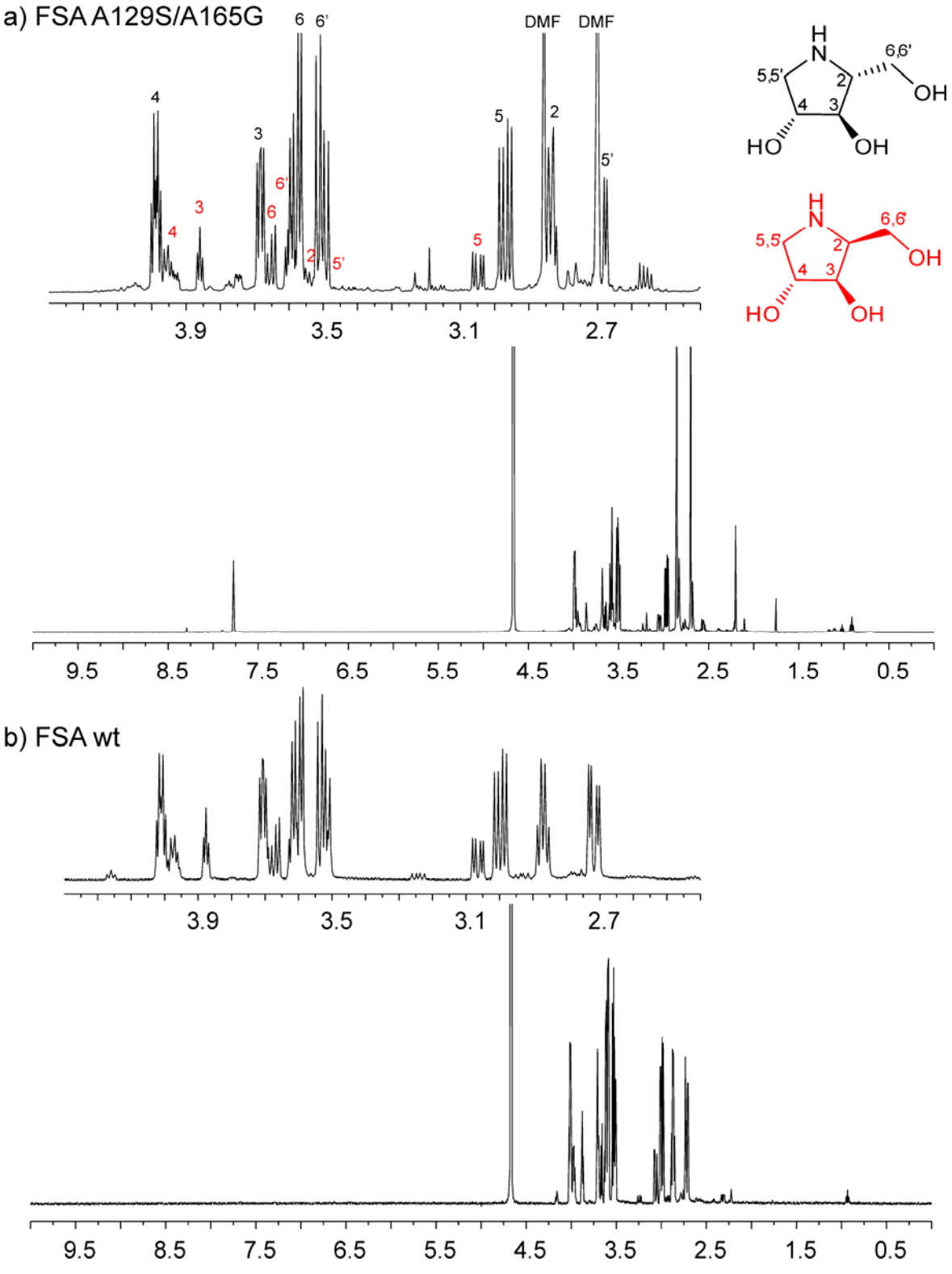


Figure S9. Observed ^{1}HNMR spectrum of 1,4-dideoxy-1,4-imino-D-arabinitol (DAB) obtained from the aldol adduct **7**. a) product from the FSA A129S/A165G double mutant. b) product from the FSA wild-type.



E) Computational Methods.

Protein-substrate models were built with the package Schrödinger Suite 2010 (Schrödinger, LLC, New York, NY, 2010) through its graphical interphase Maestro.[14] The program Macromodel[15] with its default force field OPLS 2005, a modified version of the OPLS-AA force field,[16] and GB/SA water solvation conditions[17] were used for all energy calculations. Coordinates of *E. coli* D-fructose-6-phosphate aldolase (FSA) were obtained from the Protein Data Bank[18] at Brookhaven National Laboratory (entry 1L6W). The 1L6W structure is a homodecamer formed by two packed ring-like pentamers.[19] A major interaction within each pentamer is through the C-terminal helix from one monomer (residues 199-215), which runs across the active site of the neighboring subunit. Therefore, in order to reduce the computational cost, a large fragment from one monomer (chain A, residues 1-199), containing one catalytic centre, and a small fragment of the contiguous one (chain B, residues 199-220) were used to built the protein-substrate models. The structure of the FSA A129S/A165G was generated by mutating *in silico* the wild-type protein. Protein structures were prepared using the Protein Preparation Wizard included in Maestro to remove solvent molecules and ions, adding hydrogens, setting protonation states and minimizing the energy using the OPLS force field. The hemiaminal covalent complexes of fructose-6-phosphate bound into the active centre of FSA wild-type and of (5*R*)- and (5*S*)-**5** bound into the active centre of the FSA A129S/A165G mutant were modeled by modification of a bound glyceraldehide molecule present in the crystal structure. The structures of the protein-substrate complexes were minimized, first applying constraints to the protein (force constant = 25 kcal $Å^{-2}$ mol^{-1}) to avoid large changes on its structure, and afterwards allowing free movement of the whole system. Then the conformational space of the bound hemiaminal ligands was sampled using the mixed torsional – low mode conformational search method implemented in Macromodel.[20] Finally, the lowest energy conformers found were further minimized allowing free movement of the whole system until reaching a gradient <0.01 kcal mol^{-1} $Å^{-1}$. Figures were prepared with the software package MOE.[21]

F) References

1. A. L. Concia, C. Lozano, J. A. Castillo, T. Parella, J. Joglar and P. Clapés, *Chem. Eur. J.*, 2009, **15**, 3808.
2. J. Calveras, M. Egido-Gabás, L. Gómez, J. Casas, T. Parella, J. Joglar, J. Bujons and P. Clapés, *Chem. Eur. J.*, 2009, **15**, 7310.

3. L. Espelt, T. Parella, J. Bujons, C. Solans, J. Joglar, A. Delgado and P. Clapés, *Chem. Eur. J.*, 2003, **9**, 4887.
4. M. M. Bradford, *Anal. Biochem.*, 1976, **72**, 248.
5. M. Schürmann and G. A. Sprenger, *J. Biol. Chem.*, 2001, **276**, 11055.
6. J. A. Castillo, J. Calveras, J. Casas, M. Mitjans, M. P. Vinardell, T. Parella, T. Inoue, G. A. Sprenger, J. Joglar and P. Clapés, *Org. Lett.*, 2006, **8**, 6067.
7. Z. Q. Zhang and A. G. Marshall, *J. Am. Soc. Mass Spectrom.*, 1998, **9**, 225.
8. J. A. Castillo, C. Guérard-Hélaine, M. Gutiérrez, X. Garrabou, M. Sancelme, M. Schürmann, T. Inoue, V. Hélaine, F. Charmantray, T. Gefflaut, L. Hecquet, J. Joglar, P. Clapés, G. A. Sprenger and M. Lemaire, *Adv. Synth. Catal.*, 2010, **352**, 1039.
9. L. Espelt, J. Bujons, T. Parella, J. Calveras, J. Joglar, A. Delgado and P. Clapés, *Chem. Eur. J.*, 2005, **11**, 1392.
10. M. Sugiyama, Z. Hong, P. H. Liang, S. M. Dean, L. J. Whalen, W. A. Greenberg and C.-H. Wong, *J. Am. Chem. Soc.*, 2007, **129**, 14811.
11. S. G. Lee, K. H. Park and Y.-J. Yoon, *J. Heterocycl. Chem.*, 1998, **35**, 711.
12. Y. F. Wang, D. P. Dumas and C.-H. Wong, *Tetrahedron Lett.*, 1993, **34**, 403.
13. R. J. Molyneux, Y. T. Pan, J. E. Tropea, A. D. Elbein, C. H. Lawyer, D. J. Hughes and G. W. J. Fleet, *J. Nat. Prod.*, 1993, **56**, 1356.
14. Maestro, version 9.1, Schrödinger, LLC, New York, NY, 2010
15. MacroModel, version 9.8, Schrödinger, LLC, New York, NY, 2010
16. W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225.
17. W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127.
18. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235.
19. S. Thorell, M. Schurmann, G. A. Sprenger and G. Schneider, *J. Mol. Biol.*, 2002, **319**, 161.
20. G. Chang, W. C. Guida and W. C. Still, *J. Am. Chem. Soc.*, 1989, **111**, 4379; I. Kolossváry and W. C. Guida, *J. Am. Chem. Soc.*, 1996, **118**, 5011; I. Kolossvary and W. C. Guida, *J. Comput. Chem.*, 1999, **20**, 1671.
21. Molecular Operating Environment (MOE), 2010.10; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2010