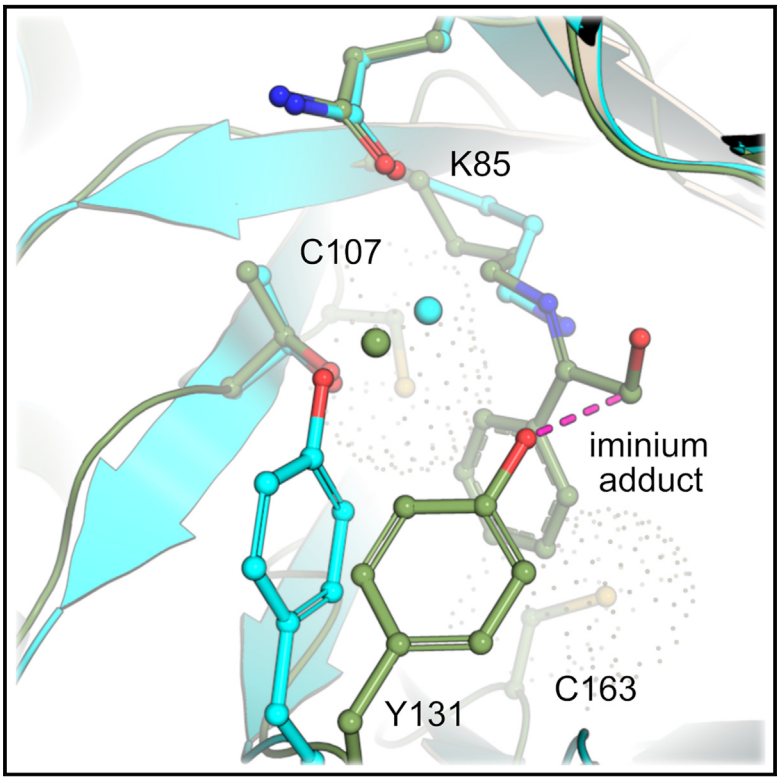
Short article

Structure

Structure of the iminium reaction intermediate in an engineered aldolase explains the carboligation activity toward arylated ketones and aldehydes

Graphical abstract



Highlights

- Cryo-EM structure of an engineered aldolase
- Captured iminium reaction intermediates of bulky ketone substrates
- Structure guided enzyme engineering for improved biocatalytic potential

Authors

Hans Hebert, Eda Sönmez, Pasi Purhonen, Mikael Widersten

Correspondence

hansh@kth.se (H.H.), mikael.widersten@kemi.uu.se (M.W.)

In brief

Hebert et al. determined the cryo-EM structure of an engineered fructose 6-phosphate aldolase. The enzyme catalyzes the carboligation between aryl substituted ketones and aldehydes. This structure contains the critical iminium reaction intermediate trapped in the active site, providing new mechanistic information.

Hebert et al., 2024, Structure *32*, 1–5
September 5, 2024 © 2024 The Author(s). Published by Elsevier Inc.
https://doi.org/10.1016/j.str.2024.06.011

CellPress

Please cite this article in press as: Hebert et al., Structure of the iminium reaction intermediate in an engineered aldolase explains the carboligation activity toward arylated ketones and aldehydes, Structure (2024), https://doi.org/10.1016/j.str.2024.06.011

Structure

CellPress
OPEN ACCESS

Short article

Structure of the iminium reaction intermediate in an engineered aldolase explains the carboligation activity toward arylated ketones and aldehydes

Hans Hebert,[1,*] Eda Sönmez,[2,3] Pasi Purhonen,[1] and Mikael Widersten[2,4,*]
[1]School of Engineering Sciences in Chemistry, Biotechnology and Health, Department of Biomedical Engineering and Health Systems, KTH Royal Institute of Technology, 14152 Huddinge, Sweden
[2]Department of Chemistry – BMC, Box 576, SE-751 23 Uppsala, Sweden
[3]Present address: Department of Chemistry – Ångström, Uppsala University, Box 256, SE-751 05, Uppsala, Sweden
[4]Lead contact
*Correspondence: hansh@kth.se (H.H.), mikael.widersten@kemi.uu.se (M.W.)
https://doi.org/10.1016/j.str.2024.06.011

SUMMARY

Two structures of fructose 6-phosphate aldolase, the wild-type and an engineered variant containing five active-site mutations, have been solved by cryoelectron microscopy (cryo-EM). The engineered variant affords production of aldols from aryl substituted ketones and aldehydes. This structure was solved to a resolution of 3.1 Å and contains the critical iminium reaction intermediate trapped in the active site. This provides new information that rationalizes the acquired substrate scope and aids in formulating hypotheses of the chemical mechanism. A Tyr residue (Y131) is positioned for a role as catalytic acid/base during the aldol reaction and the different structures demonstrate mobility of this amino acid residue. Further engineering of this fructose 6-phosphate aldolase (FSA) variant, guided by this new structure, identified additional FSA variants that display improved carboligation activities with 2-hydroxyacetophenone and phenylacetaldehyde.

INTRODUCTION

Fructose 6-phosphate aldolase (FSA) from *Escherichia coli*[1] is a type I aldolase that has been employed for biocatalytic carboligations of a range of short-chain ketones as donor substrates and aliphatic and aryl substituted aldehydes as acceptor substrates.[2–11] An advantageous feature of FSA for biocatalytic transformations, as compared to a large majority of tested aldolases, is the relaxed dependency on phosphorylated substrates and the facile recombinant production of the enzyme. A limitation of wild-type FSA, however, is the low acceptance and catalytic efficiency with bulky ketone donor substrates, although engineered enzyme variants have been isolated that accept aliphatic ketones beyond the size of dihydroxyacetone, a putative physiological substrate.[8,12]

We have recently reported the isolation of FSA variants that catalyze the aldol addition of 2-hydroxyacetophenone (**1**, Figure 1)) and ring-substituted derivatives of phenylacetaldehyde.[13] The active variants were isolated from a low-throughput screen of a library of FSA variants where the parent enzyme contained three point substitutions A129G/R134V/S166G ("VGG"), into which randomized mutagenesis was introduced at residues L107 and L163.[13] The isolates that displayed the highest activity with ketone **1** and aldehyde **2** contained a Cys residue at position 163.

A structure model of one of the most active variants $VGGC_{107}C_{163}$ was built *in silico* from a crystal structure of an R134V/S166G ("VG") double mutant of FSA (PDB: 7qxf[13]). The model underwent molecular dynamics simulation for studies of a critical reaction intermediate, but the lack of solid structural data limited the detailed interpretation of the observed structure-dynamics-activity relationships.

Here, we present the tertiary structures of histidine-tagged wild-type FSA and of the $VGGC_{107}C_{163}$ variant, determined by cryoelectron microscopy (cryo-EM) (Table 1). The latter structure contains the crucial iminium (protonated Schiff base) reaction intermediate captured in the active site (Figure 2). The structure with bound iminium was achieved from a mixture of the $VGGC_{107}C_{163}$ enzyme and ketone **1**.

RESULTS AND DISCUSSION

The structure explains the acquired substrate scope

The A129G mutation in the parent A129G/R134V/S166G ("VGG") enzyme was originally inserted to increase space in the binding site of the double R134V/S166G ("VG") variant but does not by itself improve the carboligation efficiency of **1** + **2**.[13] The additional L107C/L163C substitutions introduced in the $VGGC_{107}C_{163}$ variant widen and shape the active site cavity further. The thiol groups of C107 and C163 are both in VdW contact with the phenyl ring of bound iminium as is the Cα of G129. Together these substitutions shape the binding cavity for snug fit of the iminium intermediate (Figure 2B). The structure aids in

This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Please cite this article in press as: Hebert et al., Structure of the iminium reaction intermediate in an engineered aldolase explains the carboligation activity toward arylated ketones and aldehydes, Structure (2024), https://doi.org/10.1016/j.str.2024.06.011

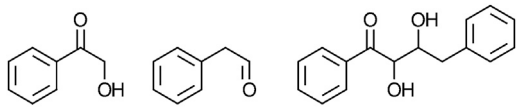


Figure 1. Compounds discussed[a]

[a]**1**, 2-hydroxyacetophenone, **2**, phenylacetaldehyde, **3**, 1,4-diphenyl-2,3-dihydroxybutan-1-one.

rationalizing the observations in catalytic activities of previously identified $VGGX_{107}C_{163}$ variants, where V107 was comparable to C107 and activities fell in the order of A107, I107, N107, G107, W107, and F107.[13] Amino acids with side chains larger than valine would interfere with the binding through sterical clashes whereas glycine would provide additional space in the binding pocket that could invite to non-productive binding of the ketone substrate and/or the iminium intermediate.

Table 1. Data collection, image processing, and refinement statistics

	Protein (FSA)	
	Wild type	$VGGC_{107}C_{163}$
Sample preparation		
Grid	Quantifoil 2.1	Quantifoil 2.1
Cryo-specimen freezing	Vitrobot 1	Vitrobot 1
Data collection		
Electron microscope	Titan KRIOS G2	Titan KRIOS G2
Detecting device	Gatan K3	Gatan K3
Total electron exposure ($e^-/Å^2$)	48.6	50.04
Dose rate ($e^-/Å^2/s$)	18.7	23.17
Number of frames in each movie	45	40
Exposure time (sec)	2.6	2.2
Defocus range (μm)	−0.6–−2.0	−0.6–−2.0
Pixel size (Å)	0.83	0.83
Image processing		
Processing software	Scipion/RELION2.1	Scipion/RELION2.1
Movie selection	9816 → 3560 → 1479	8165 → 1388
Initial particle images (no.)	251491	307612
Final particle images (no.)	162766	115055
Symmetry imposed	D5	D5
Map resolution (Å)	2.8	3.14
Refinement statistics		
Ramachandran	0/2.34/97.66	0/1.21/98.79
Clash score	18.21	5.90
CC (mask/box/peaks/volume)	0.87/0.96/0.84/0.85	0.87/0.70/0.69/0.83
Fsc map/model (Å)	3.9	3.1
Overall B-value	190	156
PDB accession code	8s7h	8s7i
EMDB accession code	EMD-19772	EMD-19773

Implications of structural dynamics as critical for catalysis and stereoselectivity

The active-site structures of wild-type FSA, determined either from X-ray crystallography[14] or by cryo-EM (the histidine-tagged version, this work) display no differences in the positions of catalytic groups. Superimposing these structures with the $VGGC_{107}C_{163}$ enzyme with the iminium intermediate bound, however, reveals specific changes that provide potentially mechanistic information. The side chain of Y131, identified as critical for the activity of FSA[15] and proposed as possibly directly involved in general acid-base catalysis, is rotated approximately 100°, as compared to its position in the structures of the wild type. This rotation shifts the position of the phenolic group by approximately 3 Å (Figures S1 and S2). In addition, the position of a water molecule, implicated as a proton relay, and participating in a hydrogen bond network with Y131, Q59, and T109, is also shifted to a similar extent (Figures 2D and S2). The observed position of Y131 places the phenol group <3 Å away from the α-carbon of bound iminium. In order for the aldol addition to proceed, deprotonation of this α-carbon is required to form the nucleophilic enamine (Scheme 1) and Y131 appears well placed for a role as a catalytic base, as judged from this structure. Furthermore, the different positions of Y131 between these structures suggest the possibility of substantial movement and dynamics of this residue.

The diastereomer selectivity of the $VGGC_{107}C_{163}$ enzyme favors synthesis of the syn isomer of aldol 3 in a ratio of syn:anti of 9:1 (with the (2*S*, 3*R*)-configuration as the major enantiomer).[13] In order for the syn diasteromer to be formed the incoming aldehyde must react with the enzyme enamine at the *si* face. Manually placing aldehyde 2 in a presumably productive pose, in the $VGGC_{107}C_{163}$ active site causes clashes with Y131 which suggests that Y131 must shift position for the reaction to proceed. Y131 placed at a position similar to that observed in the wild-type structures, however, would allow for productive aldehyde binding with the *si* face toward the enamine. This would also place the phenolic group of Y131 well positioned for protonation of the oxyanion formed on (pro-)C3 in the enzyme bound addition product (Scheme 1).

Structure guided engineering for further improved carboligation activities

The design of the enzyme library from which the $VGGC_{107}C_{163}$ variant was isolated was originally guided by the structure of the R134V/S166G double mutant. Now, with access to the current structure also providing explicit information on the positioning of the catalytically critical iminium intermediate formed with ketone **1**, a second generation library was constructed. Again, to limit the screening effort, two amino acid residues were chosen for randomized mutagenesis, F57 and A165. Both residues are placed to potentially be involved in stabilization of iminium intermediates formed with arylated ketones. Simulated structures of the carbinolamine intermediate preceding iminium formation (Scheme 1)[13] together with the current structure were used to guide the protein library design. F57 and A165 are situated on opposite sides of the iminium adduct (Figure S3) thereby lowering the risk of secondary effects from introduction of (new) unfavorable interactions introduced during the randomized mutagenesis.

Please cite this article in press as: Hebert et al., Structure of the iminium reaction intermediate in an engineered aldolase explains the carboligation activity toward arylated ketones and aldehydes, Structure (2024), https://doi.org/10.1016/j.str.2024.06.011

CellPress
OPEN ACCESS

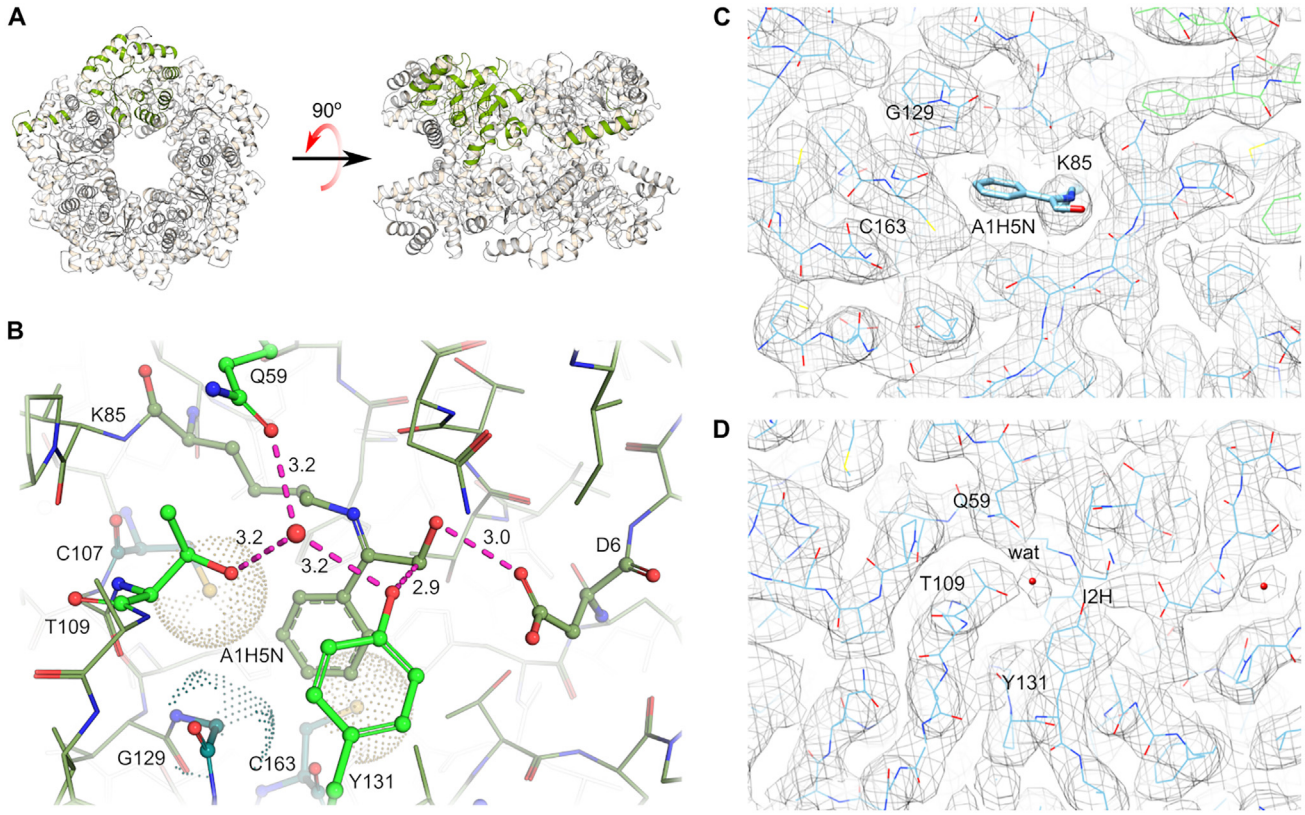


Figure 2. Structure of the VGGC$_{107}$C$_{163}$ variant of FSA determined by cryoelectron microscopy (PDB entry 8s7i)

(A) The decameric structure of FSA showing the arrangement of two stacked pentameric rings. The "A" subunit is colored in dark green. (left) Top view, (right) side view.

(B) The active site of VGGC$_{107}$C$_{163}$ (chain A) with the linked iminium intermediate formed following reaction of K85 and ketone **1** ("A1H5N").

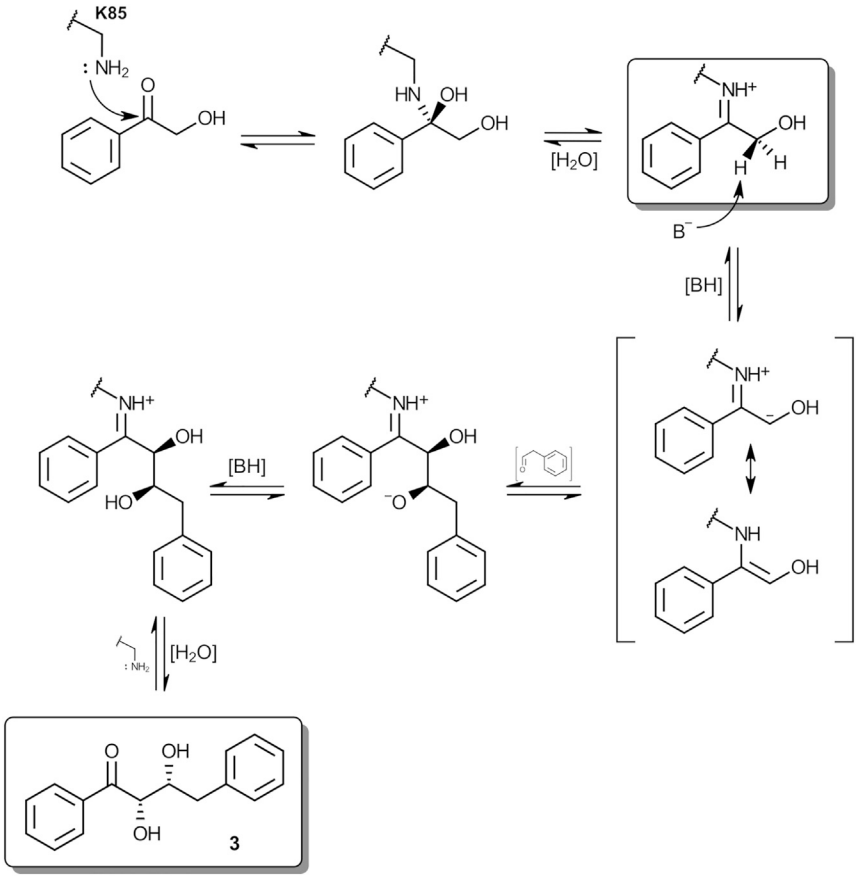
(C) VGGC$_{107}$C$_{163}$ post-processed map at 1.5σ together with the refined model highlighting the modfied K85 residue (A1H5N) and in (D), the presumably catalytic water molecule at the active site (wat). Subunit A is shown in light blue and a small part of subunit B in the upper right corner in (C) in green. Figure constructed with Pymol (A and B) and UCSF Chimera (C and D).

Out of a total of 227 library clones screened, 69 unique VGGCCX$_{57}$X$_{165}$ FSA variants were identified, cloned, and expressed in *E. coli*. 64 of these FSA variants were successfully isolated in purified form, corresponding to a 16% coverage of all possible variants (20^2) (Figure S4). Four variants, IP (where the first letter is the substitution at F57 and the latter refers to the substitution at A165), GP, KP, and KW, did not yield measurable amounts of soluble protein. The insert of a proline at position 165, which is in a loop connecting β strand 7 and α helix 7, (residues 164–170)[14] which already in the parent enzyme includes an S166G mutation, may inflict unfavorable sterical constraints possibly affecting folding negatively. Nonetheless, certain variants with an A165P mutation were purified successfully (FP, MP, NP, SP, and VP), however lower concentrations were obtained.

The catalytic abilities to produce aldol **3** from **1** and **2** were screened in reactions with fixed concentrations of substrates and enzymes and for a standardized incubation time (see the STAR methods for details). Activities were evaluated on the basis of catalyzed production of aldol **3**. Out of the 64 purified variants, 60 displayed detectable activity in the aldol reaction (Table S1; Figure S5A). It is noteworthy that the apparently relatively unstable A165P variants (NP, SP, and VP) were among the more active enzymes. The variant displaying the highest improvement in product formation, as compared to the VGGC$_{107}$C$_{163}$ parent, was VGGCCF$_{57}$I$_{165}$. This enzyme contains only an A165I point mutation which, in this assay, provides a 55% increase in product formation. As expected, the majority of mutations introduced to the active site were unbeneficial and resulted in lower carboligation activities as compared to the parent enzyme.

The most active VGGCCX$_{57}$X$_{165}$ variants showed decreased diastereomer selectivity, as compared to the parent (Table S1; Figure S5B), displaying syn:anti ratios of approximately 80:20. The underlying structure-function relationships resulting in the properties of these variants are currently speculative but certain features can be observed, (i), substitutions at position 57 appears less important for the increase in product formation. (ii), at position 165, either a proline or a residue with a β-branched side chain (Thr or Ile) is found, which may suggest a local shift in the orientation and/or a change of the dynamics of the loop. In the current structure with the iminium intermediate bound, the side chain of A165 is within VdW distance to the phenyl ring of the iminium, thus (any) substitution at that position would influence formation of a Michaelis complex as well as the

Please cite this article in press as: Hebert et al., Structure of the iminium reaction intermediate in an engineered aldolase explains the carboligation activity toward arylated ketones and aldehydes, Structure (2024), https://doi.org/10.1016/j.str.2024.06.011



Scheme 1. Model of FSA $VGGC_{107}C_{163}$ catalyzed aldol addition[a]

[a]The acid/base group ("BH/B^-") is suggested to be the Y131 phenol.

formation and stabilization of the carbinol and iminium intermediates. A bulkier side chain as in FI would sterically clash with the iminium as it is bound in the structure with $VGGC_{107}C_{163}$ which suggests that either the positioning of this loop or of the iminium must be different in a corresponding complex in the FI variant.

In conclusion, we have reported two new structures of FSA solved by cryo-EM. The wild-type structure is for all purposes indistinguishable from the available crystal structure. The structure of the laboratory engineered $VGGC_{107}C_{163}$ variant represents the first example of an aldolase enzyme that is able to catalyze aldol addition of an arylated ketone with likewise aryl substituted aldehydes such as derivatives of phenylacetaldehyde and contributes important information on the acquired substrate acceptance for 2-hydroxyacetophenone (**1**). The structure guided additional engineering of this FSA variant resulted in the isolation of variants of further improved activity with ketone 1 and phenylacetaldehyde.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Cryo electronmicrosopy
  - Enzyme engineering
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.str.2024.06.011.

ACKNOWLEDGMENTS

We thank Rasmus Kejving for assisting in production of $VGGCCX_{57}X_{165}$ proteins. This work was funded by Byggmästare Olle Engkvists Stiftelse (#218-0061). The data was collected at the Cryo-EM Swedish National Facility funded by the Knut and Alice Wallenberg, Family Erling Persson and Kempe Foundations, SciLifeLab, Stockholm University and Umeå University.

AUTHOR CONTRIBUTIONS

H.H. solved and refined the structures. E.S. performed the protein engineering, isolation, and characterization of the $VGGCCX_{57}X_{165}$ aldolase variants. P.P. collected the cryoelectron microscopy data on wild-type and the $VGGC_{107}C_{163}$ variant of FSA. M.W. conceived the project and was responsible for overall planning. All authors contributed to writing of the manuscript.

Please cite this article in press as: Hebert et al., Structure of the iminium reaction intermediate in an engineered aldolase explains the carboligation activity toward arylated ketones and aldehydes, Structure (2024), https://doi.org/10.1016/j.str.2024.06.011

CellPress
OPEN ACCESS

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 10, 2024
Revised: May 28, 2024
Accepted: June 19, 2024
Published: July 15, 2024

REFERENCES

1. Schürmann, M., and Sprenger, G.A. (2001). Fructose-6-phosphate aldolase is a novel class I aldolase from *Escherichia coli* and is related to a novel group of bacterial transaldolases. J. Biol. Chem. *276*, 11055–11061.
2. Castillo, J.A., Calveras, J., Casas, J., Mitjans, M., Vinardell, M.P., Parella, T., Inoue, T., Sprenger, G.A., Joglar, J., and Clapés, P. (2006). Fructose-6-phosphate aldolase in organic synthesis: Preparation of D-fagomine, *N*-alkylated derivatives, and preliminary biological assays. Org. Lett. *8*, 6067–6070.
3. Samland, A.K., and Sprenger, G.A. (2006). Microbial aldolases as C–C bonding enzymes—unknown treasures and new developments. Appl. Microbiol. Biotechnol. *71*, 253–264.
4. Concia, A.L., Lozano, C., Castillo, J.A., Parella, T., Joglar, J., and Clapés, P. (2009). d-Fructose-6-phosphate aldolase in organic synthesis: cascade chemical-enzymatic preparation of sugar-related polyhydroxylated compounds. Chem. Eur J. *15*, 3808–3816.
5. Clapés, P., Fessner, W.-D., Sprenger, G.A., and Samland, A.K. (2010). Recent progress in stereoselective synthesis with aldolases. Curr. Opin. Chem. Biol. *14*, 154–167.
6. Clapés, P., and Garrabou, X. (2011). Current trends in asymmetric synthesis with aldolases. Adv. Synth. Catal. *353*, 2263–2283.
7. Fesko, K., and Gruber-Khadjawi, M. (2013). Biocatalytic methods for C–C bond formation. ChemCatChem *5*, 1248–1272.
8. Güclü, D., Szekrenyi, A., Garrabou, X., Kickstein, M., Junker, S., Clapés, P., and Fessner, W.-D. (2016). Minimalist protein engineering of an aldolase provokes unprecedented substrate promiscuity. ACS Catal. *6*, 1848–1852.
9. Ma, H., Engel, S., Enugala, T.R., Al-Smadi, D., Gautier, C., and Widersten, M. (2018). New stereoselective biocatalysts for carboligation and retro-aldol cleavage reactions derived from d-fructose 6-phosphate aldolase. Biochemistry *57*, 5877–5885.
10. Al-Smadi, D., Enugala, T.R., Kessler, V., Mhashal, A.R., Lynn Kamerlin, S.C., Kihlberg, J., Norberg, T., and Widersten, M. (2019). Chemical and biochemical approaches for the synthesis of substituted dihydroxybutanones and di- and tri-hydroxypentanones. J. Org. Chem. *84*, 6982–6991.
11. Hélaine, V., Gastaldi, C., Lemaire, M., Clapés, P., and Guérard-Hélaine, C. (2022). Recent advances in the substrate selectivity of aldolases. ACS Catal. *12*, 733–761.
12. Roldan, R., Sanches-Moreno, I., Scheidt, T., Hélaine, V., Lemaire, M., Parella, T., Clapés, P., Fessner, W.-D., and Guérard-Hélaine, C. (2017). Breaking the dogma of aldolase specificity: simple aliphatic ketones and aldehydes are nucleophiles for fructose-6-phosphate aldolase. Chem. Eur J. *23*, 5005–5009.
13. Cornelius, E.C., Bartl, M., Persson, L.J., Xiong, R., Cederfelt, D., Rad, F.M., Norberg, T., Engel, S., Marklund, E.G., Dobritzsch, D., and Widersten, M. (2023). Engineered aldolases catalyzing stereoselective aldol reactions between aryl-substituted ketones and aldehydes. Catal. Sci. Technol. *13*, 4978–4987.
14. Thorell, S., Schürmann, M., Sprenger, G.A., and Schneider, G. (2002). Crystal structure of decameric fructose-6-phosphate aldolase from *Escherichia coli* reveals inter-subunit helix swapping as a structural basis for assembly differences in the transaldolase family. J. Mol. Biol. *319*, 161–171.
15. Stellmacher, L., Sandalova, T., Leptihn, S., Schneider, G., Sprenger, G.A., and Samland, A.K. (2015). Acid–base catalyst discriminates between a fructose 6-phosphate aldolase and a transaldolase. ChemCatChem *7*, 3140–3151.
16. de la Rosa-Trevín, J.M., Quintana, A., del Cano, L., Zaldívar, A., Foche, I., Gutiérrez, J., Gómez-Blanco, J., Burguet-Castell, J., Cuenca-Alba, J., Abrishami, V., et al. (2016). Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. J. Struct. Biol. *195*, 93–99.
17. Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.
18. Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 216–221.
19. de la Rosa-Trevin, J.M., Otón, J., Marabini, R., Zaldívar, A., Vargas, J., Carazo, J.M., and Sorzano, C.O.S. (2013). Xmipp 3.0: an improved software suite for image processing in electron microscopy. J. Struct. Biol. *1*, 321–328.
20. Fernandez-Leiro, R., and Scheres, S.H.W. (2017). A pipeline approach to single-particle processing in RELION. Acta Cryst. D Struct. Biol. *73*, 496–502.
21. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera – a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.
22. Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Cryst. D Struct. Biol. *74*, 531–544.
23. Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Cryst. D Biol. Cryst. *60*, 2126–2132.
24. Moriarty, N.W., Grosse-Kunstleve, R.W., and Adams, P.D. (2009). electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. Acta Cryst. D Biol. Cryst. *65*, 1074–1080.
25. Dale, G.E., Schönfeld, H.-J., Langen, H., and Stieger, M. (1994). Increased solubility of trimethoprim-resistant type SI DHFR from *Staphylococcus aureus* in *Escherichia coli* cells overproducing the chaperonins GroEL and GroES. Protein Eng. *7*, 925–931.
26. Hamnevik, E., Blikstad, C., Norrehed, S., and Widersten, M. (2014). Kinetic characterization of *Rhodococcus ruber* DSM 44541 alcohol dehydrogenase A. J. Mol. Catal. B Enzymol. *99*, 68–78.

Please cite this article in press as: Hebert et al., Structure of the iminium reaction intermediate in an engineered aldolase explains the carboligation activity toward arylated ketones and aldehydes, Structure (2024), https://doi.org/10.1016/j.str.2024.06.011

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Escherichia coli BL21-AI	Invitrogen	Cat# 11540146
Escherichia coli XL1-Blue	Agilent	Cat# 200228
Chemicals, peptides, and recombinant proteins		
bicine	ThermoScientific	Cat# A14957.30
2-hydroxyacetophenone	Merck	Cat# 445134
Bacto tryptone	ThermoFisher Scientific	Cat# 211705
Bacto yeast extract	ThermoFisher Scientific	Cat# 212720
NaCl	Merck	Cat# S7653
ampicillin	ThermoFisher Scientific	Cat# J63807.09
kanamycin sulfate	ThermoFisher Scientific	Cat# 450811000
L-arabinose	Merck	Cat# A3256
isopropyl β-*d*-thiogalctopyranoside	ThermoFisher Scientific	Cat# 34060
imidazole	Merck	Cat# I2399
lysozyme	ThermoFisher Scientific	Cat# 89833
DNase I	Invitrogen	Cat# 18047019
Deposited data		
Atomic coordinates *E. coli* fructose 6-phosphate aldolase (C-terminally 5-His tagged)	This paper	8s7h
Maps of *E. coli* fructose 6-phosphate aldolase (C-terminally 5-His tagged)	This paper	EMD-19772
Atomic coordinates *E. coli* fructose 6-phosphate aldolase L107C/A129G/R134V/L163C/S166G (C-terminally 5-His tagged)	This paper	8s7i
Maps of *E. coli* fructose 6-phosphate aldolase L107C/A129G/R134V/L163C/S166G (C-terminally 5-His tagged)	This paper	EMD-19773
Oligonucleotides		
See Table S2	This paper.	N/A
Recombinant DNA		
pREP4-groESL	Dale et al.[25]	N/A
pGT7ADHAA-5H	Hamnevik et al.[26]	N/A
Software and algorithms		
Scipion	de la Rosa-Trevín et al.[16]	https://scipion.i2pc.es/
MotionCor2	Zheng et al.[17]	https://emcore.ucsf.edu/ucsf-software
CTFFIND4	Rohou & Grigorieff[18]	https://grigoriefflab.umassmed.edu/ctffind4
Xmipp	de la Rosa-Trevín et al.[19]	http://xmipp.i2pc.es/
Relion	Fernandez-Leiro et al.[20]	https://www3.mrc-lmb.cam.ac.uk/relion
UCSF-Chimera	Pettersen et al.[21]	https://www.cgl.ucsf.edu/chimera/
Phenix	Afonine et al.[22]	https://phenix-online.org/
Coot	Emsley & Cowtan[23]	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
eLBOW	Moriarty et al.[24]	https://phenix-online.org/documentation/reference/elbow.html
Phenix.douse	N/A	https://phenix-online.org/version_docs/1.20.1-4487/reference/douse.html

Please cite this article in press as: Hebert et al., Structure of the iminium reaction intermediate in an engineered aldolase explains the carboligation activity toward arylated ketones and aldehydes, Structure (2024), https://doi.org/10.1016/j.str.2024.06.011

CellPress
OPEN ACCESS

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Mikael Widersten (mikael.widersten@kemi.uu.se).

Materials availability

Expression clones generated in this study are available from the lead contact.

Data and code availability

- Cryo-EM maps and the coordinates of the structures have been deposited at the EMDB and PDB and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Bacterial host cells: *Escherichia coli* XL1-Blue and BL21-AI were cultured in 2TY liquid medium, at 30°C, in rotary shaker incubators, or in Luria Bertani (LB) medium, at 37°C, when cultured on agar plates. Media were fortified with appropriate antibiotics dependent on carried plasmids as described in the method details section.

METHOD DETAILS

Cryo electronmicrosopy

Data on grid preparations, data collection and refinement statistics are summarized in Table 1.

Grid preparation and data collection

FSA (wild type, C-terminally 5-His tagged) in 40 mM bicine, pH 8.35 at a concentration of 11.5 mg/mL was diluted with 40 mM bicine, pH 8.35 buffer to 2.3 mg/mL. FSA $VGGC_{107}C_{163}$ in 40 mM bicine, 2 mM 2-hydroxyacetophenone (**1**, Figure 1) pH 8 at a concentration of 10 mg/mL was diluted to 2 mg/mL 3.0 μL of the diluted samples were placed to glow-discharged (1 min, 36 mA) Quantifoil R2/1 200 mesh Cu-grids. The sample was incubated on the grid for 5 s and blotted for 3 s with Whatman 595 blotting paper at 100% humidity, 16°C, with a Vitrobot Mark I and plunge frozen into liquid ethane. The grids were transferred to liquid nitrogen for storage.

A dataset of pure FSA (wild type) and FSA $VGGC_{107}C_{163}$ in the presence of **1** were collected with a Gatan K3 camera on Titan Krios G2 microscope equipped with an energy filter set to 20 eV. In both sessions data were collected at 300 kV with a defocus range from −0.6 μm to −2.0 μm in counting mode using EPU software in faster acquisition mode.

The FSA dataset yielded 9816 micrographs (105,000× magnification, pixel size = 0.8617 Å, total dose 48.64 $e^-/Å^2$, 45 frames, resulting in 1.08 $e^-/Å^2$ dose/frame). The FSA $VGGC_{107}C_{163}$ mutant dataset gave 8165 micrographs (105,000× magnification, pixel size = 0.8566 Å, total dose 50.04 $e^-/Å^2$, 40 frames, resulting in 1.251 $e^-/Å^2$ dose/frame).

Processing of cryoEM data and model building

Movies imported into the Scipion platform[16] were subjected to motion correction using MotionCor2.[17] Obvious useless and suboptimal recordings were discarded followed by ctf correction using CTFFIND4.[18] Based on the quality of Thon rings (defocus less than 3 μm, no visible astigmatism and isotropic rings to high resolution) the image datasets were reduced further to 1479 elements for wild type and 1388 for the $VGGC_{107}C_{163}$ mutant, respectively. Particle picking was carried out using the manual/autopicking procedure in Xmipp[19] from 2× binned and contrast-inverted micrographs (Figure S6A). Intermediate stacks of particles were extracted from undose weighted images at a box size of 160 × 160 pixels of a size corresponding to 0.86 Å on the specimen level. At a later stage in processing the sampling was corrected to 0.83 Å based on a correlation distribution against known molecular models. 2D-classification identified non-particle extractions (Figure S6B) which were discarded. 3D initial model, auto-refinement and classification in Relion[20] using standard parameter settings were repeated until homogeneous datasets were obtained (162,766 particles for the wild type and 115,055 for the mutant). CTF refinement, further auto-refinement and post-processing completed processing of the datasets into maps with D5 symmetry. The resolution of the maps (Figure S6D) were obtained from pairs of unfiltered half-maps using the EMDB FSC validation server (https://www.ebi.ac.uk/pdbe/emdb/validation/fsc/).UCSF-Chimera[21] was utilized for initial docking of molecular models previously obtained from X-ray crystallography (pdb entries 1lw6[14] and 7qxf[13]). A flow scheme of the processing of the data is shown in Figure S6E. The models were further refined in real space using Phenix.[22] Coot[23] was used at some stages of model building. For the wild-type enzyme the 1lw6 structure could be used without any further modifications since the sequence is identical to the one in this study. For the $VGGC_{107}C_{163}$ mutant, however, the corresponding residues were changed. Furthermore during model refinement it was obvious that the map showed density not fully accounted for by the model. In particular a continuous density at the expected active site could be fitted with substrate **1** added to the protein solution. Thus, it was expected that the experiment had captured the iminium reaction intermediate (Figure 2C). The resulting modified lysine residue at position 85 was not available as a chemical component in the pdb database. Restraints for this compound, now termed A1H5N, could be obtained from the

Please cite this article in press as: Hebert et al., Structure of the iminium reaction intermediate in an engineered aldolase explains the carboligation activity toward arylated ketones and aldehydes, Structure (2024), https://doi.org/10.1016/j.str.2024.06.011

eLBOW routine[24] in Phenix enabling further refinement against the map. Water molecules within hydrogen bonding distance to the protein and corresponding to significant round-shaped peaks in the maps were localized using DOUSE in Phenix (Figure 2D). Final refinements generated models having high ranks with regard to standard geometry and favorable correspondence between maps and models (Table 1; Figure S7).

Enzyme engineering

Buffers and growth media

All bacteria were grown in 2TY medium (16 g/L tryptone, 10 g/L yeast extract and 5 g/L NaCl). *Escherichia coli* XL1-blue cells (Agilent) were plated on LB-Amp plates (15 g/L bacto-agar, 10 g/L tryptone, 5 g/L yeast extract and 10 g/L NaCl, 100 μg/mL ampicillin), whereas *E. coli* BL21-AI[pREP4-groELS][25] cells (Invitrogen) were plated on LB-AmpKan plates (100 μg/mL ampicillin and 30 μg/mL kanamycin). A prepared 4% (w/v) L-arabinose stock and 1 mM isopropyl β-D-thiogalactoside (IPTG) were used for induction of protein expression.

For protein purification 40 mM bicine buffer (pH 8.35) was used with varying salt concentrations, except the final desalting buffer which contained no additional salts. Binding, washing, and elution buffers contained 0.5 M NaCl with 20 mM, 40 mM, and 300 mM imidazole, respectively. All bacterial pellets were resuspended in lysis buffer (binding buffer with 0.1 mg/mL lysozyme, 10 μg/mL DNaseI and 1× Thermo Scientific Halt Protease Inhibitor Cocktail, EDTA-Free (100×) (Cat. No: 78425)).

Construction of the VGGCCX$_{57}$X$_{165}$ library

Library construction was achieved applying a stepwise PCR where the randomly mutated FSA variant genes were assembled from individual DNA-fragments (Figure S8). Oligonucleotide sequences of PCR primers are provided in Table S2. Every gene fragment was a mixture of amplified DNA from eight parallel reactions: 50 μL each; 10 μL 5×Phusion HF buffer, 1 μL dNTPs (10 mM), 5 μL forward primer (10 μM), 5 μL reverse primer (10 μM), 5 ng of template DNA (VGGC$_{107}$C$_{163}$),[S8] 1.5 μL dimethyl sulfoxide (DMSO), 0.5 μL Thermo Scientific Phusion High-Fidelity DNA polymerase and sterile MilliQ-water. The forward oligonucleotide primers containing the mutated codons NDT, VMA, ATG and TGG were added in 12:6:1:1 M ratios to a total volume of 5 μL – specifically 3, 1.5, 0.25 and 0.25 μL, respectively. All reactions were performed in gradient PCR using the (thermal settings in Table S3). Fragment 1 was produced from oligonucleotides pUC-Fwd and VGGCCF57-Rev while VGGCC-F57-Fwd- (mixture of -NDT, -VMA, -ATG and -TGG) and VGGCC-A165-Rev were used to amplify fragment 2 and VGGCC-A165-Fwd- (mixture of -NDT, -VMA, -ATG and -TGG) and pUC-Rev for fragment 3. Fragment 4 was obtained from an equimolar mixture of fragments 1 and 2 as template DNA together with primers pUC-Fwd and VGGCC-A165-Rev. Finally, the completely assembled gene library was obtained from a mixture of fragments 3 and 4 as template DNA and pUC-Fwd and pUC-Rev as primers.

The gene library (~3 μg DNA) was subjected to digestion with *Xho*I and *Bcu*I overnight at room temperature. The digested gene library fragments were agarose gel purified and ligated overnight at room temperature to purified plasmid vector digested with the same enzymes at a molar ratio vector:fragment ~1:5 (pGT7ADHA-5H).[26] The ligation reactions were heated to 70°C and subsequently transformed into *E. coli* XL1-Blue cells. A 14-mL culture (2TY with 100 μg/mL ampicillin) of transformed cells was incubated at 37°C with shaking (200 rpm) overnight followed by preparation of the plasmid content encoding the FSA VGGCCX$_{57}$X$_{165}$ gene library. Obtained plasmids were sent for sequencing to confirm mutagenesis and estimate sequence diversity.

Purified library plasmid DNA was transformed into chemically competent *E. coli* BL21-AI[pREP4-groELS] and plated on LB-AmpKan plates. Transformation of library DNA was repeated every 1–2 weeks to obtain fresh cells for the analysis.

Protein production

Proteins were expressed, purified and analyzed, essentially as described before.[13] Purity of proteins was estimated by SDS-PAGE stained with Coomassie Brilliant Blue R-250 (Figure S4).

Activity screening

The supernatant of stopped reactions with different FSA VGGCCX$_{57}$X$_{165}$ mutants were analyzed by reversed phase HPLC (C-18 column, Ascentis, 25 × 0.46 cm, 5 μm bead size) and an LC-20AD pump using the Shimadzu Providence system. A gradient elution of 40–90% methanol was used with 50 mM sodium phosphate, pH 3.0, at a flow-rate of 0.5 mL/min with 50 min measurement time. A sample (20 μL) was injected with an autosampler (SIL-20A) and eluted components were detected at 212 nm to detect the internal standards, benzyl alcohol, or at 244 nm for the aldol products, using an SPD-M20A diode array detector. Product amounts were quantified from a standard curve established previously (605,759 mA×min×nmol^{-1}).[13]

Triplicate reactions of purified protein were set-up in final concentrations of 10 μM (2 nmol) enzyme, 20 mM (4 μmol) 2-hydroxyacetophenone (**1**, Figure 1) and 5 mM (1 μmol) phenylacetaldehyde (**2**) in a total volume of 200 μL with 2.5% (v/v) acetonitrile. Blank reactions without enzyme were run in parallel. Reactions were incubated at 30°C for 22 h in the dark. The reactions were stopped by addition of methanol containing 2 mM (0.4 μmol) benzyl alcohol to a final concentration of 50% (v/v). Protein was removed by centrifugation at 17,000×*g* for 30 min before transferring 100 μL to HPLC-vials for analysis. The product, 2,3-dihydroxy-1,4-diphenylbutanone (**3**) peaks were detected at 244 nm with approximate retention times of 36.5 min and 38.5 for the *anti*- and *syn*-diasteromers, respectively. The diastereomeric *anti*:*syn* (dr) ratios of produced **3** were determined from integrated peak areas following reversed phase separation as described above. Peaks were identified by comparing the retention times and spectral properties of formed products to those of previously synthesized *anti*-**3** together with previously run *anti*- and *syn*-mixtures of **3**.[10]

Please cite this article in press as: Hebert et al., Structure of the iminium reaction intermediate in an engineered aldolase explains the carboligation activity toward arylated ketones and aldehydes, Structure (2024), https://doi.org/10.1016/j.str.2024.06.011

CellPress
OPEN ACCESS

QUANTIFICATION AND STATISTICAL ANALYSIS

Cryo-EM data collection and refinement statistics are summarized in Table 1.