Supporting Information: An Enzyme Cascade Synthesis of Vanillin

Tobias Klaus [1], Alexander Seifert [1], Tim Häbe [1], Jürgen Pleiss [1], Bettina M. Nestl [1] and Bernhard Hauer [1,*]

[1] Institute of Biochemistry and Technical Biochemistry, Department of Technical Biochemistry, Universitaet Stuttgart, Allmandring 31, 70569 Stuttgart, Germany, tobias.klaus@curevac.com (T.K.), seifert@lfb.fr (A.S.), tim.haebe@ernaehrung.uni-giessen.de (T.H.), Juergen.pleiss@itb.uni-stuttgart.de (J.P.), bettina.nestl@itb.uni-stuttgart.de (B.M.N.)

* Correspondence: bernhard.hauer@itb.uni-stuttgart.de; Tel.:+49-711-685-64523

Table S1. Initial screening results for the conversion of 3-methylanisole **(1)** catalysed by all 24 variants plus wild type of the CYP102A1 minimal mutant library. Lysate of empty vector expression (pET-22b(+), pET-28a(+)) was used as a negative control and showed no conversion.

Amino Acid at Position		**87**				
		A	**F**	**I**	**L**	**V**
328	**A**	++	+ [a]	+	+	+++
	F	++	++	++	+	+++
	I	++	++	++	++	+++
	L	+	+++	++	–	+++
	V	+	++	++	–	+++

–, no conversion; +, conversion < 5%; ++, conversion < 25%; +++, conversion > 25%; [a] CYP102A1 wild type.

Table S2. Results of the conversion of 3-methylanisole (**1**) with all investigated variants of CYP102A1.

CYP102A1 Variant	**Product Yield (% of Initial Substrate)**					**Total Conversion (%)**
	2	**3**	**6**	**9**	**10**	
wild type	–	–	<1	–	–	<1
A328L	<1	2.7 ± 0.2	26.6 ± 2.5	<1	–	31.0 ± 2.8
F87V	<1	<1	12.5 ± 0.2	<1	–	14.9 ± 0.3
F87V/A328F	2.2 ± 0.2	1.2 ± 0.1	16.5 ± 1.3	7.1 ± 0.4	–	26.9 ± 2.0
F87V/A328I	6.2 ± 0.5	<1	5.6 ± 0.3	4.9 ± 0.3	–	16.9 ± 1.1
F87V/A328L	9.1 ± 0.2	1.6 ± 0.1	23.2 ± 0.4	12.0 ± 0.8	13.5 ± 0.8	59.3 ± 1.8
F87V/A328V	4.9 ± 0.5	<1	23.1 ± 0.9	3.0 ± 0.5	2.1 ± 0.1	34.1 ± 2.0
F87V/A328L/ L437A	1.6 ± 0.1	<1	3.1 ± 0.2	4.7 ± 0.2	–	9.5 ± 0.5
F87V/A328L/ L437F	1.7 ± 0.2	<1	5.7 ± 0.9	7.3 ± 1.0	<1	15.5 ± 2.3
F87V/A328L/ L437I	20.0 ± 0.9	<1	14.0 ± 0.7	7.4 ± 1.3	19.3 ± 0.5	61.3 ± 2.9
F87V/A328L/ L437V	22.3 ± 2.0	<1	10.5 ± 1.0	7.0 ± 0.8	18.5 ± 2.7	58.9 ± 5.3
R47L/Y51F	<1	–	1.5 ± 0.1	–	–	1.6 ± 0.1
R47L/Y51F/ A328L	<1	1.2 ± 0.1	12.5 ± 1.3	<1	–	14.8 ± 1.5

R47L/Y51F/F87V	2.1 ± 0.1	1.6 ± 0.1	35.1 ± 3.1	1.2 ± 0.1	2.0 ± 0.3	42.5 ± 3.7
R47L/Y51F/F87V/A328F	2.1 ± 0.3	1.3 ± 0.2	26.1 ± 3.5	5.1 ± 0.4	5.8 ± 1.4	40.3 ± 5.8
R47L/Y51F/F87V/A328I	10.4 ± 1.1	<1	9.5 ± 0.8	6.9 ± 0.5	1.6 ± 0.3	28.8 ± 2.7
R47L/Y51F/F87V/A328L	4.9 ± 0.5	<1	13.9 ± 0.9	9.7 ± 0.3	7.9 ± 0.9	37.4 ± 2.6
R47L/Y51F/F87V/A328V	5.0 ± 0.3	<1	24.0 ± 0.8	3.0 ± 0.1	2.1 ± 0.2	34.9 ± 1.4
R47L/Y51F/F87V/A328L/L437I	13.5 ± 0.4	<1	10.1 ± 0.5	7.7 ± 0.2	17.8 ± 0.8	50.4 ± 1.8

–, not detected or <0.05%; **2**, 3-methoxybenzyl alcohol; **3**, 4-methylguaiacol; **6**, 4-methoxy-2-methylphenol; **9**, *m*-cresol; **10**, methylhydroquinone. Samples were analysed by GC-FID. Negative controls showed no conversion.

Table S3. Results of the conversion of 4-methylguaiacol (**3**) with R47L/Y51F-variants of CYP102A1.

CYP102A1 Variant	Conversion to Product 4 (%)
R47L/Y51F	–
R47L/Y51F/A328L	–
R47L/Y51F/F87V	<1
R47L/Y51F/F87V/A328F	1.6 ± 0.1
R47L/Y51F/F87V/A328I	3.0 ± 0.3
R47L/Y51F/F87V/A328L	<1
R47L/Y51F/F87V/A328V	4.5 ± 0.3
R47L/Y51F/F87V/A328L/L437I	3.1 ± 0.6

–, not detected or <0.05%. Samples were analysed by HPLC. Negative controls showed no conversion.

Table S4. Results of the in vivo conversion of 3-methylanisole (**1**) with a combination of the CYP102A1 variant A328L and the vanillyl alcohol oxidase (VAO) variant F454Y.

Conversion Time (h)	Product Yield (% of Initial Substrate)					Total Conversion (%)
	2	**3**	**5**	**6**	**9**	
0	–	–	–	–	–	–
0.25	–	–	–	<0.5	–	<0.5
0.5	–	–	–	<0.5	–	<0.5
1	–	–	–	<0.5	–	<0.5
2	–	<0.5	–	0.9 ± 0.1	–	1.0 ± 0.1
3	<0.5	<0.5	<0.5	3.6 ± 0.4	<0.5	4.3 ± 0.4
4	<0.5	<0.5	0.5	6.4 ± 0.1	<0.5	7.6 ± 0.2
6	<0.5	<0.5	0.8	8.1 ± 0.2	<0.5	9.7 ± 0.3
8	<0.5	<0.5	0.9 ± 0.1	9.2 ± 0.8	<0.5	10.8 ± 1.0
12	<0.5	–	1.1 ± 0.1	9.9 ± 1.3	<0.5	11.7 ± 1.4

24	<0.5	–	1.1 ± 0.3	9.9 ± 2.4	<0.5	11.7 ± 2.6

–, not detected or <0.05%; **2**, 3-methoxybenzyl alcohol; **3**, 4-methylguaiacol; **5**, vanillin; **6**, 4-methoxy-2-methylphenol; **9**, *m*-cresol. Samples were analysed by GC-FID. Negative controls showed no conversion.

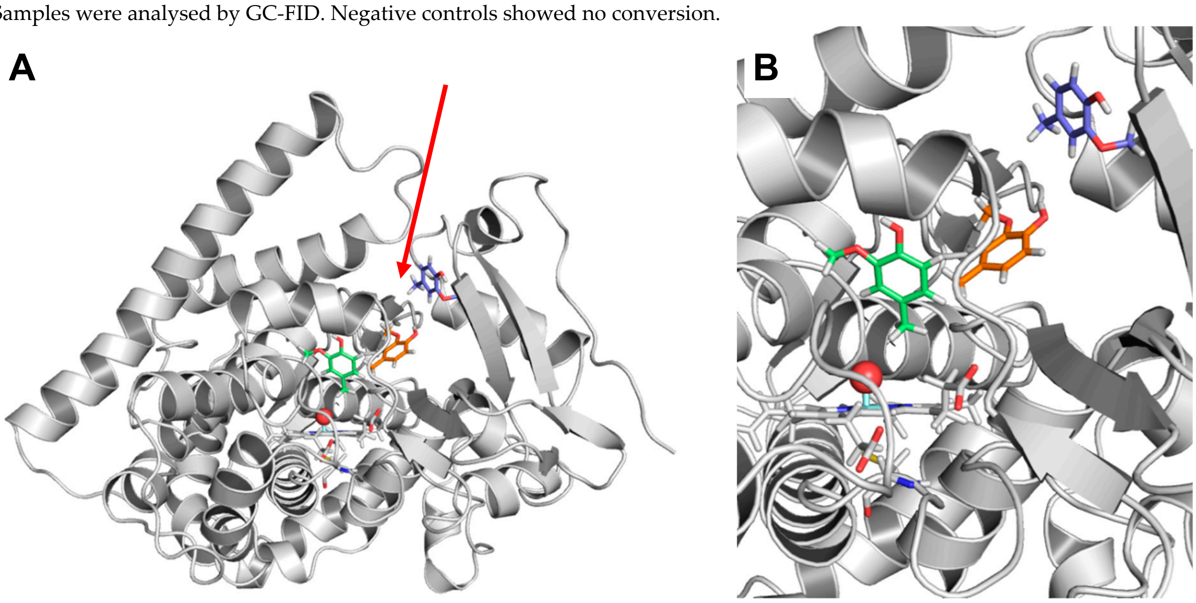

Figure S1. Regions in CYP102A1 variant F87V/A328I active site cavity where 4-methylguaiacol (**3**) was frequently observed during molecular dynamics (MD) simulations. The orange molecule shows the starting position, green—the region close to the haem center and blue—the pocket formed by the β1 sheet and the A′ helix. The substrate access channel is indicated by a red arrow. (**B**) is zoomed in on (**A**).

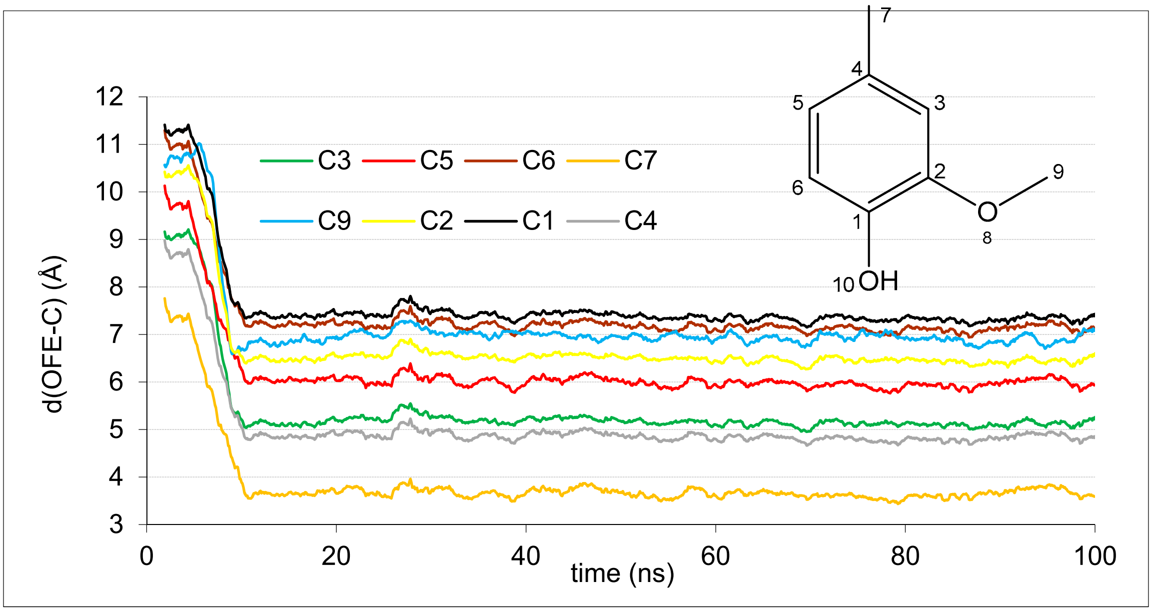

Figure S2. Distance of 4-methylguaiacol **(3)** carbon atoms from the activated oxygen of the haem in a 100 ns MD run. The molecule approaches the haem oxygen to a distance < 4 Å (C7) where it stays stable for 90 ns. Lines show a moving average of 20 data points.

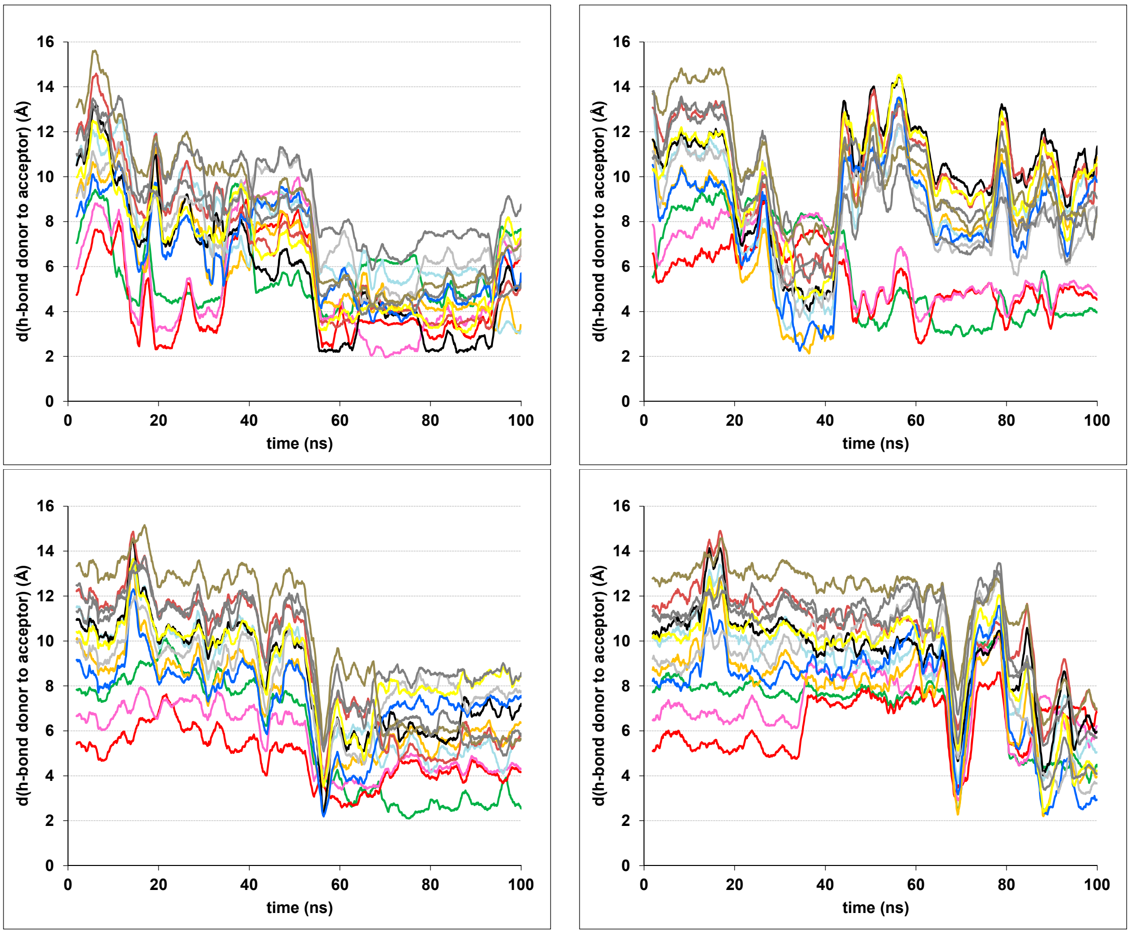

Figure S3. Representation of four exemplary MD simulations from a pool of 15 MD simulations. The distance between all possible hydrogen bond donors and acceptors of 4-methylguaiacol **(3)** and the side chain of Y51 (green, red and magenta lines) and R47 (all remaining lines) are shown. A hydrogen bond is considered present at a hydrogen bond donor to acceptor distance ≤3.5 Å. Lines show a moving average of 20 data points.

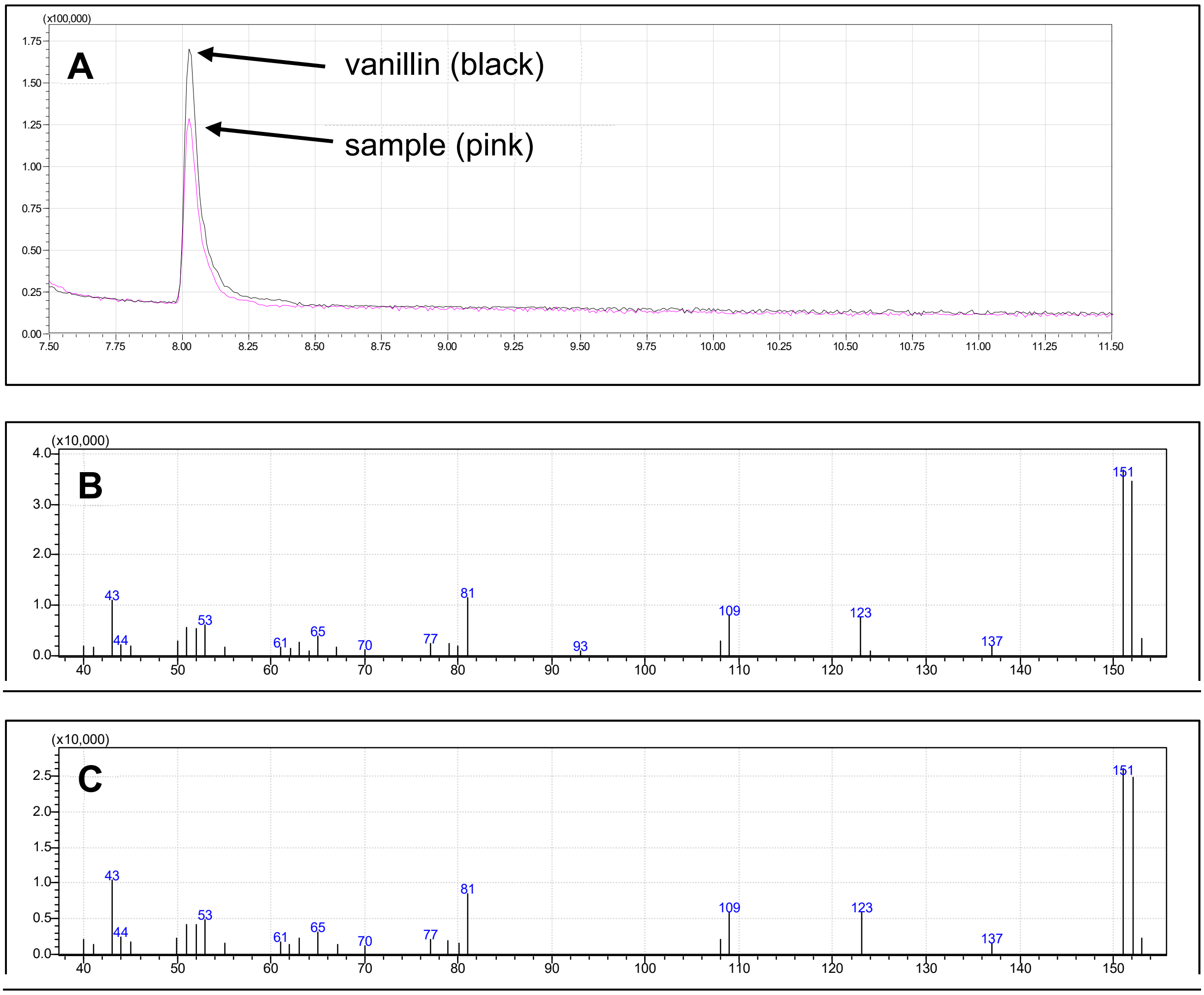

Figure S4. GC–MS analysis chromatogram (**A**) and fragmentation patterns of reference compound (**B**) (0.05 mM vanillin) and sample (**C**) (after 12 h conversion time) of the in vivo biotransformations.