

# **Chalcone-based Selective Inhibitors of a C<sub>4</sub> Plant Key Enzyme as Novel Potential Herbicides**

## **SUPPLEMENTARY INFORMATION**

G.T.T. Nguyen<sup>1</sup>, G. Erlenkamp<sup>2</sup>, O. Jäck<sup>3</sup>, A. Küberl<sup>4</sup>, M. Bott<sup>4</sup>, F. Fiorani<sup>3</sup>, H. Gohlke<sup>2</sup>, and G. Groth<sup>1\*</sup>

<sup>1</sup>Biochemical Plant Physiology and <sup>2</sup>Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf and Bioeconomy Science Center (BioSC)

<sup>3</sup>Institute of Bio- and Geosciences, IBG-2: Plant Sciences and <sup>4</sup>Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich and Bioeconomy Science Center (BioSC)

\* Correspondence and requests for materials should be addressed to G. G.:

Heinrich Heine University, Biochemical Plant Physiology, Universitätsstr. 1, 40225 Düsseldorf, Germany, Phone: +49-211-81-12822, Fax: +49-211-81-13569, Email: [Georg.Groth@hhu.de](mailto:Georg.Groth@hhu.de)

## Supplemental Methods

### *Protein structure preparation*

We retrieved the crystal structures of PEPC of the C<sub>4</sub> plant *F. trinervia* (PDB ID 3ZGE) and the C<sub>3</sub> plant *F. pringlei* (PDB ID 3ZGB)<sup>1</sup> from the Protein Data Bank<sup>2</sup>. Both structures were preprocessed with the Protein Preparation Wizard<sup>3-7</sup> of the Schrödinger suite. Bond orders were assigned, hydrogens were added, the H-bond network was optimized, and missing side chains were detected and added using Prime<sup>8,9</sup>. Finally, the systems were energy minimized using the OPLS 2005 force field resulting in a root mean square deviation (RMSD) of 0.3 Å with respect to the crystal structure<sup>10-12</sup>. Since the allosteric binding site is not affected by other monomers, we only used one monomer for further analysis<sup>13,14</sup>.

### *Ligand preparation*

The structures of the twelve ligands were sketched with ChemDraw 14<sup>15</sup>. 3D structures were generated with the LigPrep<sup>16</sup> module of the Schrödinger suite. We calculated the pKa values of all compounds with the Epik module<sup>4,17,18</sup>. For the *para* hydroxyl groups in ring A (see Table 1), a mean pKa value of  $7.9 \pm 0.8$  was computed. This value is in line with experimental values reported for derivatives<sup>19</sup>.

### *Molecular dynamics simulations*

MD simulations were carried out for **12** inside the allosteric feedback inhibitor site of C<sub>4</sub> PEPC (PDB ID 3ZGE) and **11** inside of C<sub>3</sub> PEPC (PDB ID 3ZGB). The preparation of the ligands and the setup of the simulations were performed with programs from the Ambertools15 package<sup>20</sup>. Antechamber<sup>21</sup> was used to parameterize the compounds. Atomic charges were computed with

the AM1-BCC method<sup>22,23</sup>. The ff99SB force field<sup>24</sup> was used for the protein, and the General Amber Force Field (GAFF)<sup>21</sup> was used for the compounds. All co-crystallized components, such as water, 1,2-ethandiol, sulfate, and the inhibiting aspartate were removed from both used crystal structures prior to the following preparation of the complexes. LEaP was used to generate complex models from the ligand and receptor structures. After neutralizing the systems by adding 19 or 20 Na<sup>+</sup> ions according to the charge of the ligand, the systems were placed inside a truncated octahedral periodic box of TIP3P water<sup>25</sup> with a distance between the box edges and the closest atom of the complex of at least 10 Å. The size of the systems was ~92,000 atoms. Three independent MD simulations for each complex were performed with the AMBER14 suite using the GPU-accelerated pmemd program<sup>20,26</sup>. Initially, the solvated systems were energy minimized by 500 steps of steepest descent, followed by 500 steps of conjugate gradient minimization. The non-bonded cutoff was set to 8.0 Å, and the positions of the complex atoms were restrained with a harmonic potential using a force constant of 2.0 kcal mol<sup>-1</sup> Å<sup>-2</sup>. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm<sup>27</sup>. After this step, the temperature was increased from 0 K to 300 K during 50 ps of NVT (canonical ensemble) MD simulations. The density of the solvent was adjusted by MD simulations of 50 ps length in the NPT (isothermal-isobaric) ensemble. The final equilibration step was performed for 500 ps during which the force constant for the harmonic restraints was set to zero. For all MD simulations, a time step of 2 fs was used. Production simulations were set up with a coupling time for the heat bath of 2.0 ps. The production runs achieved lengths of 200 ns.

### *AFMoC (Adaption of Fields for Molecular Comparison) analysis*

One AFMoC model each was derived for the ligands within C<sub>4</sub> and C<sub>3</sub> PEPC (PDB IDs 3ZGE and 3ZGB), respectively. The ligand superimpositions were obtained by docking with GLIDE and subsequent minimization with the MAB force field as described above. DrugScore<sup>28</sup> was used to compute receptor-based potential fields<sup>29</sup> for four atom types present within the set of ligands (atom-type notation according to Sybyl<sup>30</sup>: C.2, C.ar, O.3,O.2). The grid box was set to a size such that all ligands were embedded with a margin of at least 4 Å, and the grid spacing was set to 1 Å. The half width of the Gaussian to compute interaction fields from the potential fields (see ref.<sup>29</sup> for more details) was set to 0.85 Å. Leave-one-out (LOO) analyses were performed. Aromatic carbons and hydroxyl oxygens of the ligands contribute most to the explanation of the variance in the binding affinities of the ligands (57% and ≥37 %, respectively). Due to the limited size of the data set further leave-multiple-out crossvalidations were not performed. Additionally, LOO-analyses after random scrambling of the biological data (“y-scrambling”) were performed. They resulted in  $q^2 < 0$  for both C<sub>4</sub> and C<sub>3</sub> PEPC, indicating that, with the original biological data, no chance correlation is given.

### *Experimental design and assessments of in planta experiments*

We tested eleven compounds in a first experiment. In a follow-up experiment we tested selected compounds that showed growth inhibitory effects in the first experiment and one additional chalcone. Each treatment was repeated 12 times, except for oilseed rape in the second experiment (n=10). Pots were arranged in the growth chamber according to a fully randomized design. Initial growth of the second leaf of *A. retroflexus* plants was monitored by acquiring images of each plant from above from day zero to day two after treatment. Images were taken

from a fixed distance at 90° view angle using the camera Canon EOS 70D equipped with a Canon EF 14 mm/2.8 L II USM lens. Leaf area of single leaves was measured using the ImageJ software<sup>31</sup>, by drawing the leaf contours using a drawing tablet (Cintiq 22HD, Wacom Europe GmbH, Krefeld, Germany). Conversion of pixel to cm was achieved by using a target of known size inside the image. Six days after the treatment, half of the plants were harvested. Leaf area was assessed using a leaf area meter (LI3100C Area Meter, LI-Cor, Lincoln, Nebraska, USA) and samples were oven-dried at 65°C for seven days to obtain dry weight. Thirteen days after the treatment, this procedure was repeated with the second half of plants.

#### *Chlorophyll fluorescence measurements*

For assessment of photosynthetic parameters, *A. retroflexus* plants were grown under the abovementioned conditions until the two-leaves growth stage. 6 mm leaf discs were cut out of the second leaf and immediately placed into 96-well-plates filled with 300 µL 0.125 % Hoagland's solution with 1 mM MES-KOH (MES-monohydrate, Merck, Darmstadt, Germany) buffered to a pH of 6.5. The treatment solution contained 1.6 % (v/v) DMSO, 0.1 % (v/v) Tween<sup>®</sup>20 and 2',3',4',3,4-Pentahydroxychalcone at a concentration of 2 mM. The control treatment contained 1.6 % (v/v) DMSO, 0.1 % (v/v) Tween<sup>®</sup>20. Each leaf disc was treated with 1.5 µL solution placed on top of the leaf discs. Plates were sealed with Parafilm and placed into a climate cabinet with 12h/12h day/night-length and a temperature of 22°C at a light intensity of 130 µmol m<sup>-2</sup> s<sup>-1</sup>. 96 hours after treatment, chlorophyll fluorescence measurements were performed on leaf discs using an Imaging PAM M-Series fluorescence camera (Heinz Walz GmbH, Effeltrich, Germany). Leaf discs were dark adapted for 30 minutes, minimum and maximum chlorophyll fluorescence were obtained and leaf discs were illuminated for 30 minutes

with a light intensity of  $130 \mu\text{mol m}^{-2} \text{s}^{-1}$  PAR. Light curves were measured by stepwise PAR increments (0, 1, 21, 56, 111, 281, 396, 611, 801, 1251  $\mu\text{mol m}^{-2} \text{s}^{-1}$  PAR). Leaf discs were illuminated at each intensity level for 30s before measuring fluorescence and maximum fluorescence by applying a saturation pulse of  $2700 \mu\text{mol m}^{-2} \text{s}^{-1}$  PAR<sup>32</sup>.

#### *Hyperspectral imaging of *A. retroflexus* plants*

Images were taken with the hyperspectral line-scanner Pica XC (Resonon, MT; USA) equipped with Schneider Xenoplan lens with 23 mm focal length (Jos. Schneider Optische Werke GmbH, Bad Kreuznach, Germany). The line-scanner is integrated in a system for automatic transport and positioning of sensors in a greenhouse<sup>33</sup> and is placed inside a closed box to prevent light from outside disturbing the measurements. The box is equipped with two 400 Watt halogen lamps providing measurement light. Images were taken with an integration time of 50 ms, a framerate of  $20 \text{s}^{-1}$ , a scanning speed of  $4 \text{mm s}^{-1}$  and a spectral resolution of 4.37 nm. Plants were cultivated as in the other experiments, except that temperature and light conditions differed. Temperature conditions in the greenhouse were  $19^{\circ}\text{C}/17^{\circ}\text{C}$  at a day/night length of 16h/8h and a light intensity of approximately  $85 \mu\text{E PAR m}^{-2} \text{s}^{-2}$ . Relative humidity was approximately 60 %. Plants were treated in the three-leaves stage and treatment solutions were prepared as mentioned before. Each treatment had six replicates. Before recording hyperspectral images, dark current was measured by covering the objective lens. Digital numbers of recorded hyperspectral images were transformed into reflectance by setting them in relation to a calibrated reflectance target (Spectralon, SpehereOptics, Uhldingen, Germany) that was placed inside the image, after the dark current spectrum was subtracted from both reflectance target and sample spectra. Mean spectra of the 3<sup>rd</sup> leaf of *A. retroflexus* were extracted as well as the corresponding spectra of the

reflection target. Photochemical reflection index (PRI) was calculated according to Gamon et al.<sup>34</sup> and anthocyanin reflectance index (ARI) was calculated according to Gitelson et al.<sup>35</sup>. For false-color images, data cubes were segmented by using a threshold based on the normalized difference vegetation index (NDVI) in combination with a NIR-threshold at 830 nm to separate plant pixels from background. ARI and PRI were calculated separately for each pixel. Image processing was performed using R Software for Statistical Computing version 3.1.1<sup>36</sup>. The package ‘hyperSpec’<sup>37</sup> was used to read data files in bil-format.

#### *Statistical Analysis of in planta experiments*

All analysis were performed using R Software for Statistical Computing version 3.1.1<sup>36</sup>. Leaf area and biomass data were first analysed separately for each experiment, because treatments differed. To test whether effects of okanin (**12**) and (**10**) on *A. retroflexus* were consistent over both experiments, the combined data of both experiments were analyzed including ‘experiment’ as a categorical two-level factor. Analyses of variance were performed with  $\alpha=0.05$ . Shapiro-Wilk tests were performed to test for normal distribution of residuals, and homogeneity of residuals was checked visually. Comparisons of means were conducted using Tukey’s honest significance test using the package ‘agricolae’<sup>38</sup>. Analyses of variances were performed on spectral indices ( $\alpha=0.05$ ). Fischer’s LSD was calculated to compare treatment means using the package ‘agricolae’<sup>38</sup>. Analysis of leaf growth data was performed using the R package ‘nlme’<sup>39</sup>. A two-parametric exponential growth model was fitted to the data with compound symmetry structure including random effects for each parameter accounting for between-plant variability. We run the model including treatment effects as covariate for both parameters. However, covariate effects were not significant for the intercept parameter, whereas it was for the rate

parameter. Therefore we included the treatment effect only for the rate parameter. Gas exchange measurements were used to estimate maximal PEP carboxylation rate ( $V_{\text{pmax}}$ ) of *A. retroflexus* according to the model of von Caemmerer<sup>40</sup> at intercellular CO<sub>2</sub> partial pressure smaller 70  $\mu\text{bar}$ . A full model, with  $V_{\text{pmax}}$  varying for measurement dates, and a reduced model, with one common  $V_{\text{pmax}}$  for measurement dates, was fitted separately for each treatment and compared via F-test at  $\alpha=0.05$ . Mitochondrial respiration losses were not accounted for in the final models, because they occurred to have no significant effects in model-fitting. For data of *B. napus* we fitted the Farquhar-von Caemmerer-Berry model<sup>41</sup>. This model estimates maximal rubisco carboxylation rate ( $V_{\text{cmax}}$ ), maximal electron transport rate ( $J_{\text{max}}$ ) and mitochondrial respiration ( $R_{\text{d}}$ ). The full model was fitted with separate parameters for measurement dates. The model was stepwise reduced to common parameters for measurement dates starting with the least significant  $R_{\text{d}}$ , followed by  $J_{\text{max}}$  and  $V_{\text{cmax}}$ . Reduced models were compared to the full model via an F-test at  $\alpha=0.05$ . Constants were taken from Sharkey et al.<sup>42</sup>.

### *Bacterial cultivation conditions*

*C. glutamicum* and *P. putida* were cultivated at 30°C, *E. coli* and *B. subtilis* at 37°C. All strains were cultivated in BHI complex medium (Difco Laboratories, Detroit, USA) supplemented with 2% (w/v) glucose and in appropriate minimal media containing 2% (w/v) glucose as carbon and energy source. The minimal media used were CGXII medium<sup>43</sup> for *C. glutamicum*, M9 medium<sup>44</sup> for *E. coli* and *P. putida*, and TS (Spizizen)<sup>45</sup> medium for *B. subtilis*. Deviating from the standard composition, TS medium was supplemented with M9 trace element solution<sup>44</sup> and 20 mg/L L-tryptophan, whereas no yeast extract was added. All strains were inoculated in three biological replicates in 5 ml BHI medium containing 2% glucose and cultivated for 6 hours at 170 rpm. Then, 285  $\mu\text{l}$  of the pre-cultures were washed and inoculated into 20 ml of complex



medium or the appropriate minimal medium and shaken at 140 rpm overnight. For inoculation of the main culture, cells equivalent to an optical density at 600 nm of 0.5 (*C. glutamicum*, *E. coli*, and *P. putida*) or 0.25 (*B. subtilis*) in 800  $\mu$ L medium were washed and resuspended in fresh media. Inhibitor substances were dissolved in DMSO and added to the media to a final concentration of the  $IC_{50}$  of the *F. trinervia* PEPC and 10-fold higher and lower concentrations. Since the least polar *trans*-chalcone and 2'-hydroxychalcone precipitated in the growth media at the highest concentrations, the maximal concentration of these compounds was reduced to 400  $\mu$ M and 100  $\mu$ M, respectively. The final backscatter at 620 nm, the growth rate ( $h^{-1}$ ), and the duration of the lag phase of all cultures were determined in a BioLector microscale cultivation system in comparison to negative controls supplemented with the solvent DMSO. A longer lag phase was only assumed if the growth rates were comparable.

## Supplemental Results

### *AFMoC (Adaption of Fields for Molecular Comparison) analysis*

To substantiate the qualitative structure-activity relationship (SAR) for the chalcone compounds described in the main text, a quantitative SAR model using the protein-based AFMoC analysis was established<sup>29</sup>.

For the ligand superimposition in C<sub>4</sub>, a significant and good AFMoC model with a  $q^2$  value of 0.63 for a leave-one-out cross-validation were obtained (Table S1 and Fig. S4a). Due to the limited range of IC<sub>50</sub> values of the chalcones binding to C<sub>3</sub> PEPC (Table 1), no AFMoC model can be derived for this case<sup>46</sup>. STDEV\*COEFF contour plots for hydroxyl oxygen elucidate regions where the presence of hydroxyl oxygens will enhance or reduce binding (Fig. S4b, c). Prominent contours indicating enhanced binding are visible in the vicinity of R641, K829, and N964 as well as close to L680 and R687. These results corroborate the importance of hydroxyl groups 2', 3', 4' in ring A and 3, 4 in ring B for okanin's (**12**) high affinity. In turn, the absence of hydroxyl groups at positions 3' and 4' of ring A and positions 3 and 4 of ring B in compound **5** explains its low affinity in both PEPC variants. Finally, a contour indicating reduced binding points towards G884, confirming that hydroxyl groups at positions 5' (in **5**) or 6' (in **11**) of ring A hamper binding to C<sub>4</sub> PEPC (Table 1).

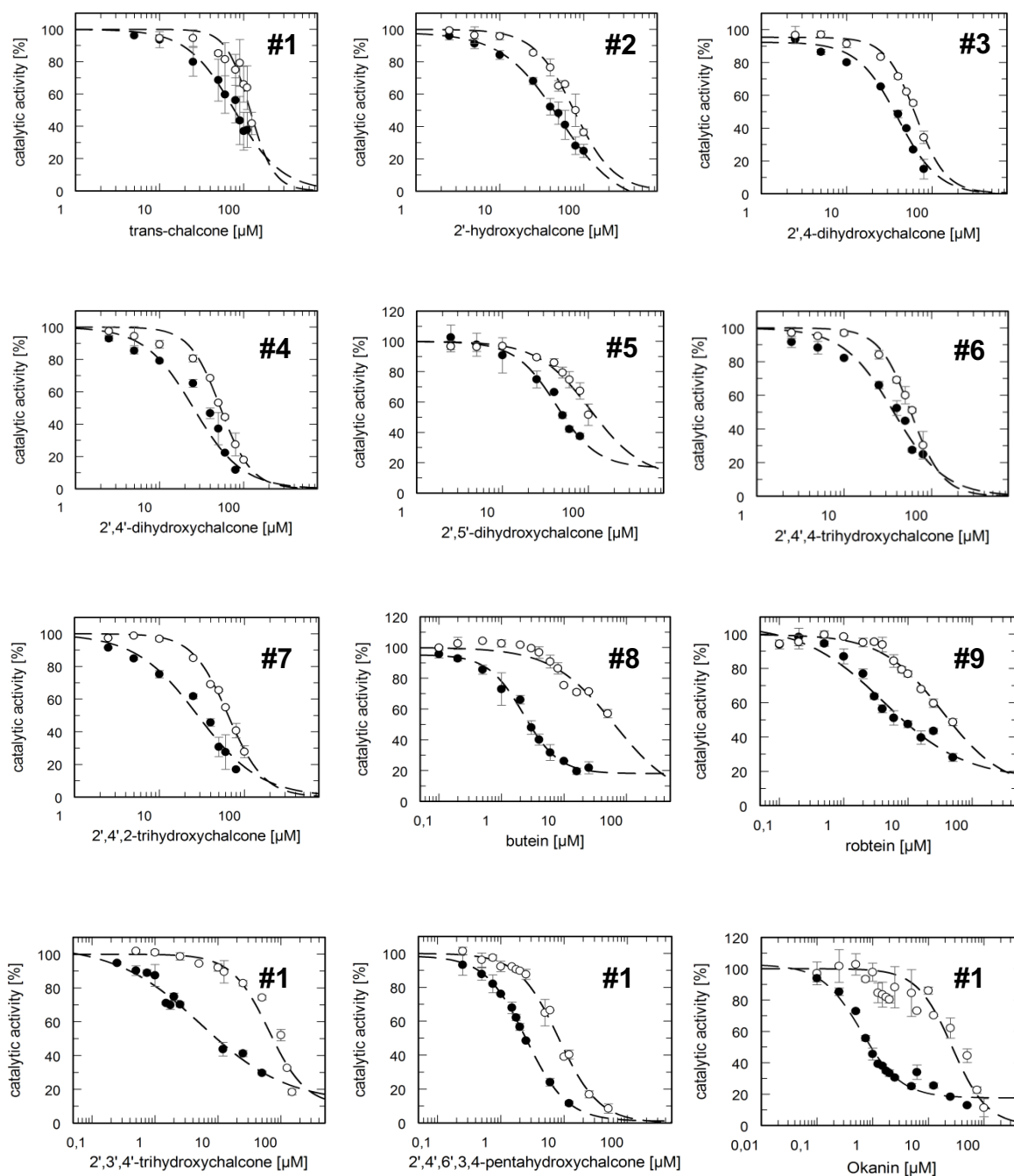
## Supplemental References

- 1 Paulus, J. K., Niehus, C. & Groth, G. Evolution of C4 phosphoenolpyruvate carboxylase: enhanced feedback inhibitor tolerance is determined by a single residue. *Mol Plant* **6**, 1996-1999 (2013).
- 2 Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Research* **28**, 235-242 (2000).
- 3 Schrödinger Suite 2014-1, Protein Preparation Wizard (Schrödinger, LLC, New York, 2014).
- 4 Schrödinger Suite 2014-1, Epik v. 2.7 (Schrödinger, LLC, New York, NY, 2013).
- 5 Schrödinger Suite 2014-1, Impact v. 6.2 (Schrödinger, LLC, New York, NY, 2014).
- 6 Schrödinger Suite 2014-1, Prime v. 2.5 (Schrödinger, LLC, New York, NY, 2014).
- 7 Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of computer-aided molecular design* **27**, 221-234 (2013).
- 8 Jacobson, M. P., Friesner, R. A., Xiang, Z. & Honig, B. On the Role of the Crystal Environment in Determining Protein Side-chain Conformations. *Journal of Molecular Biology* **320**, 597-608 (2002).
- 9 Jacobson, M. P. *et al.* A hierarchical approach to all-atom protein loop prediction. *Proteins* **55**, 351-367 (2004).
- 10 Banks, J. L. *et al.* Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comput. Chem.* **26**, 1752-1780 (2005).
- 11 Jorgensen, W. L. & Tirado-Rives, J. The OPLS Potential Functions for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *JACS* **110**, 10 (1988).
- 12 Gilson, M. K., Given, J. A., Bush, B. L. & McCammon, J. A. The Statistical-Thermodynamic Basis for Computation of Binding Affinities: A Critical Review. *Biophys. J.* **72**, 1047-1069 (1997).
- 13 Matsumura, H. *et al.* Crystal Structures of C4 Form Maize and Quaternary Complex of E. coli Phosphoenolpyruvate Carboxylases. *Structure* **10**, 1721-1730 (2002).
- 14 Muramatsu, M., Suzuki, R., Yamazaki, T. & Miyao, M. Comparison of plant-type phosphoenolpyruvate carboxylases from rice: identification of two plant-specific regulatory regions of the allosteric enzyme. *Plant Cell Physiol* **56**, 468-480 (2015).
- 15 ChemDraw v. 14.0 (PerkinElmer, 2014).
- 16 Schrödinger Suite 2014-1, LigPrep (Schrödinger, LLC, New York, NY, 2014).
- 17 Greenwood, J., Calkins, D., Sullivan, A. & Shelley, J. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *Journal of computer-aided molecular design* **24**, 591-604 (2010).
- 18 Shelley, J. C. *et al.* Epik: a software program for pK( a ) prediction and protonation state generation for drug-like molecules. *Journal of computer-aided molecular design* **21**, 681-691 (2007).
- 19 Rastelli, G., Antolini, L., Benvenuti, S. & Costantino, L. Structural bases for the inhibition of aldose reductase by phenolic compounds. *Bioorganic & Medicinal Chemistry* **8**, 1151-1158 (2000).
- 20 Amber14 (University of California, San Francisco, 2015).

- 21 Wang, J., Wolf, M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and Testing of a General Amber Force Field. *Journal of computational chemistry* **25**, 17 (2004).
- 22 Jakalian, A., Bush, B. L., Jack, D. B. & Bayly, C. I. Fast, Efficient, Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method. *Journal of computational chemistry* **21**, 15 (2000).
- 23 Jakalian, A., Jack, D. B. & Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **23**, 1623-1641 (2002).
- 24 Wang, J., Cieplak, P. & Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules. *Journal of computational chemistry* **21**, 26 (2000).
- 25 Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, 926 (1983).
- 26 Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *Journal of chemical theory and computation* **9**, 3878-3888 (2013).
- 27 Miyamoto, S. & Kollman, P. A. SETTLE: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *Journal of computational chemistry* **13**, 10 (1992).
- 28 Gohlke, H., Hendlich, M. & Klebe, G. Knowledge-based Scoring Function to Predict Protein-Ligand Interactions. *J. Mol. Biol* **295**, 20 (2000).
- 29 Gohlke, H. & Klebe, G. DrugScore Meets CoMFA: Adaption of Fields for Molecular Comparison (AFMoC) or How to Tailor Knowledge-BASED pair-Potentials to a Particular Protein. *Journal of medicinal chemistry* **45**, 17 (2002).
- 30 SYBYL-X v. 1.2 (Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144 USA, 2010).
- 31 Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat Meth* **9**, 671-675 (2012).
- 32 Maxwell, K. & Johnson, G. N. Chlorophyll fluorescence—a practical guide. *Journal of Experimental Botany* **51**, 659-668 (2000).
- 33 Fanourakis, D. *et al.* Rapid determination of leaf area and plant height by using light curtain arrays in four species with contrasting shoot architecture. *Plant Methods* **10**, 9-9 (2014).
- 34 Gamon, J. A., Serrano, L. & Surfus, J. S. The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. *Oecologia* **112**, 492-501 (1997).
- 35 Gitelson, A. A., Merzlyak, M. N. & Chivkunova, O. B. Optical Properties and Nondestructive Estimation of Anthocyanin Content in Plant Leaves. *Photochemistry and Photobiology* **74**, 38-45 (2001).
- 36 Team, R. C. R: A language and environment for statistical computing. **R Foundation for Statistical Computing, Vienna, Austria** (2014).
- 37 Beleites, C. & Sergio, V. hyperSpec: a package to handle hyperspectral data sets in R. **R package version 0.98-20150304** (2015).

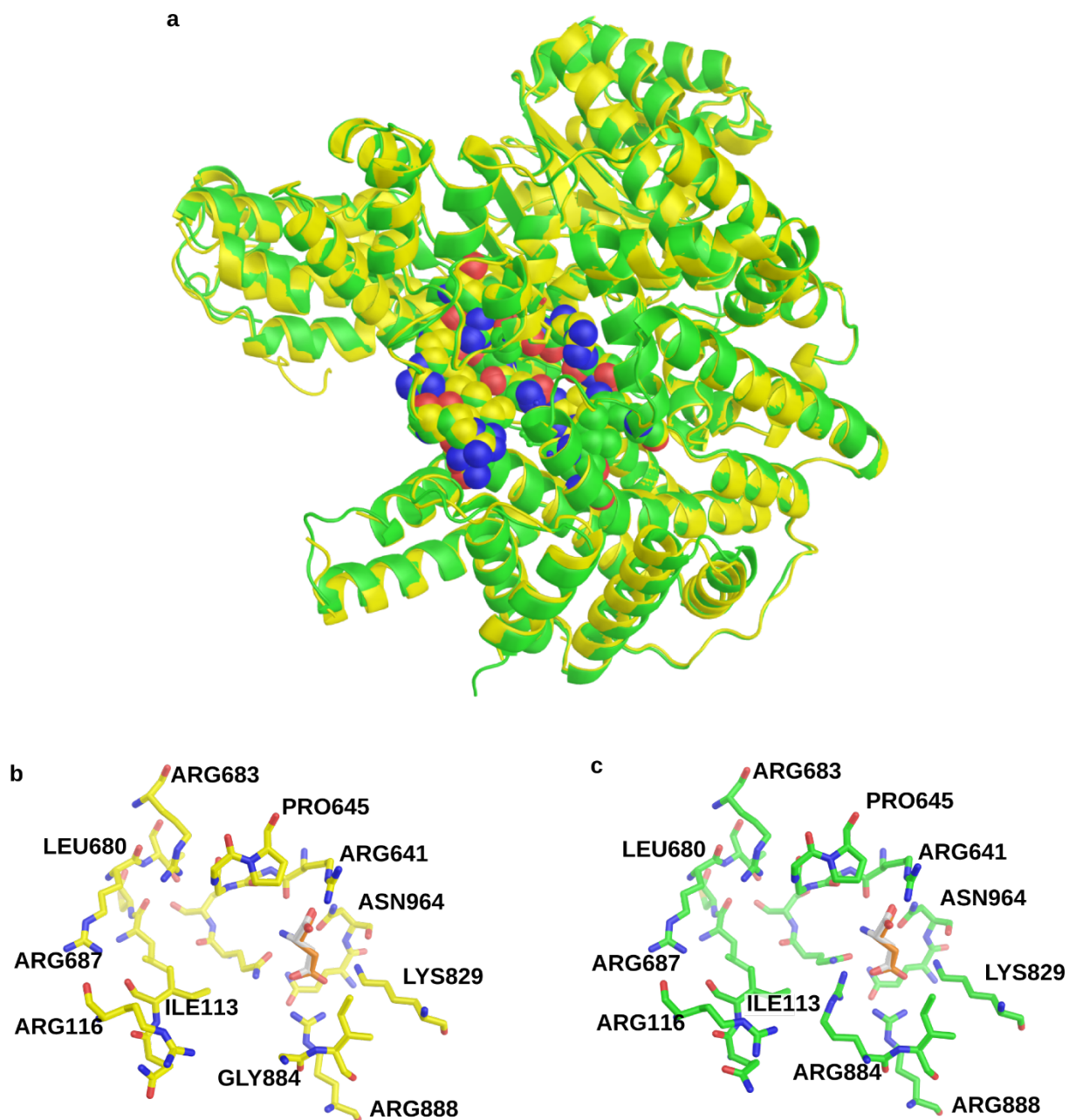
- 38 Mendiburu, F. d. agricolae: Statistical Procedures for Agricultural Research. **R package version 1.2-1** (2014).
- 39 Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & Team, R. C. nlme: Linear and Nonlinear Mixed Effects Models. **R package version 3.1-120** (2015).
- 40 von Caemmerer, S. *Biochemical models of Leaf Photosynthesis. Techniques in Plant Sciences No 2.*, 96 (Collingwood: CSIRO Publishing, 2000).
- 41 Farquhar, G. D., von Caemmerer, S. & Berry, J. A. A biochemical model of photosynthetic CO<sub>2</sub> assimilation in leaves of C<sub>3</sub> species. *Planta* **149**, 78-90 (1980).
- 42 Sharkey, T. D., Bernacchi, C. J., Farquhar, G. D. & Singsaas, E. L. Fitting photosynthetic carbon dioxide response curves for C<sub>3</sub> leaves. *Plant, Cell & Environment* **30**, 1035-1040 (2007).
- 43 Frunzke, J., Engels, V., Hasenbein, S., Gätgens, C. & Bott, M. Co-ordinated regulation of gluconate catabolism and glucose uptake in *Corynebacterium glutamicum* by two functionally equivalent transcriptional regulators, GntR1 and GntR2. *Molecular microbiology* **67**, 305-322 (2008).
- 44 Lee, D. H. & Palsson, B. O. Adaptive Evolution of *Escherichia coli* K-12 MG1655 during Growth on a Nonnative Carbon Source, L-1,2-Propanediol. *Applied and environmental microbiology* **76**, 4158-4168 (2010).
- 45 Hauser, P. M. & Karamata, D. A rapid and simple method for *Bacillus subtilis* transformation on solid media. *Microbiology* **140**, 1613-1617 (1994).
- 46 Thibaut, U. *et al.* in *3D QSAR in Drug Design. Theory, Methods and Applications* (ed H. Kubinyi) 711-716 (ESCOM, 1993).

## Supplemental Figures



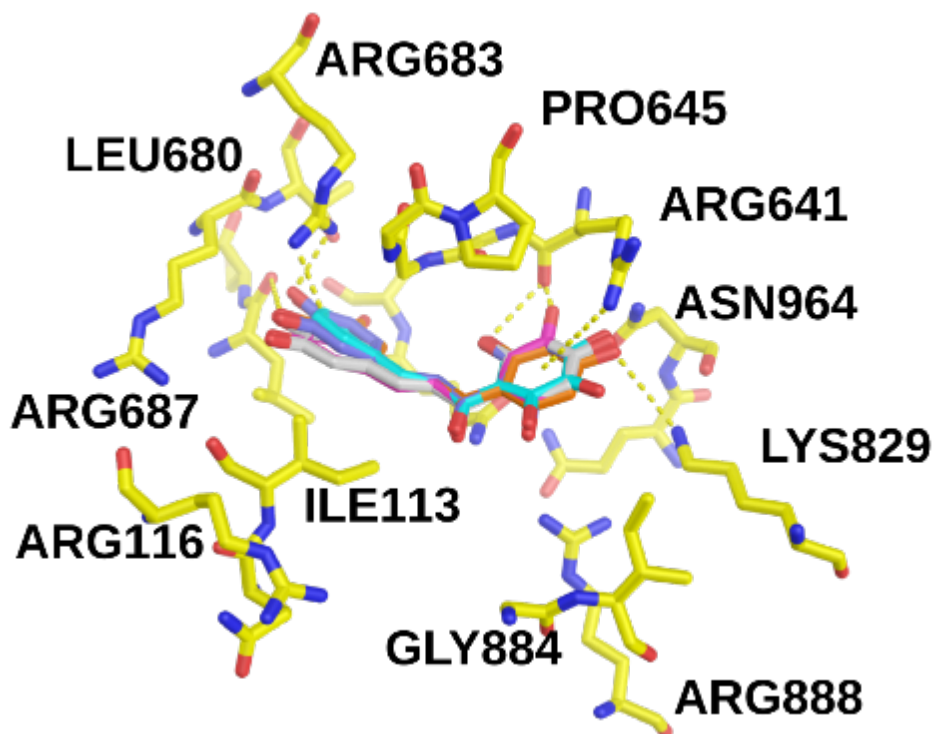
**Figure S1: IC<sub>50</sub> determination of 12 chalcones on PEPCs from *F. trinervia* and *F. pringlei* using activity assay.**

Inhibition curves are fitted and the IC<sub>50</sub> values are calculated for the C<sub>4</sub> (●) and C<sub>3</sub> (○) PEPC. Okanin (**12**) with the best potency and selectivity on C<sub>4</sub> over C<sub>3</sub> has the largest shift between two fitting curves. Error bars represent standard errors of two independent measurements.



**Figure S2: Redocking of Asp in C<sub>4</sub> and C<sub>3</sub> PEPC.**

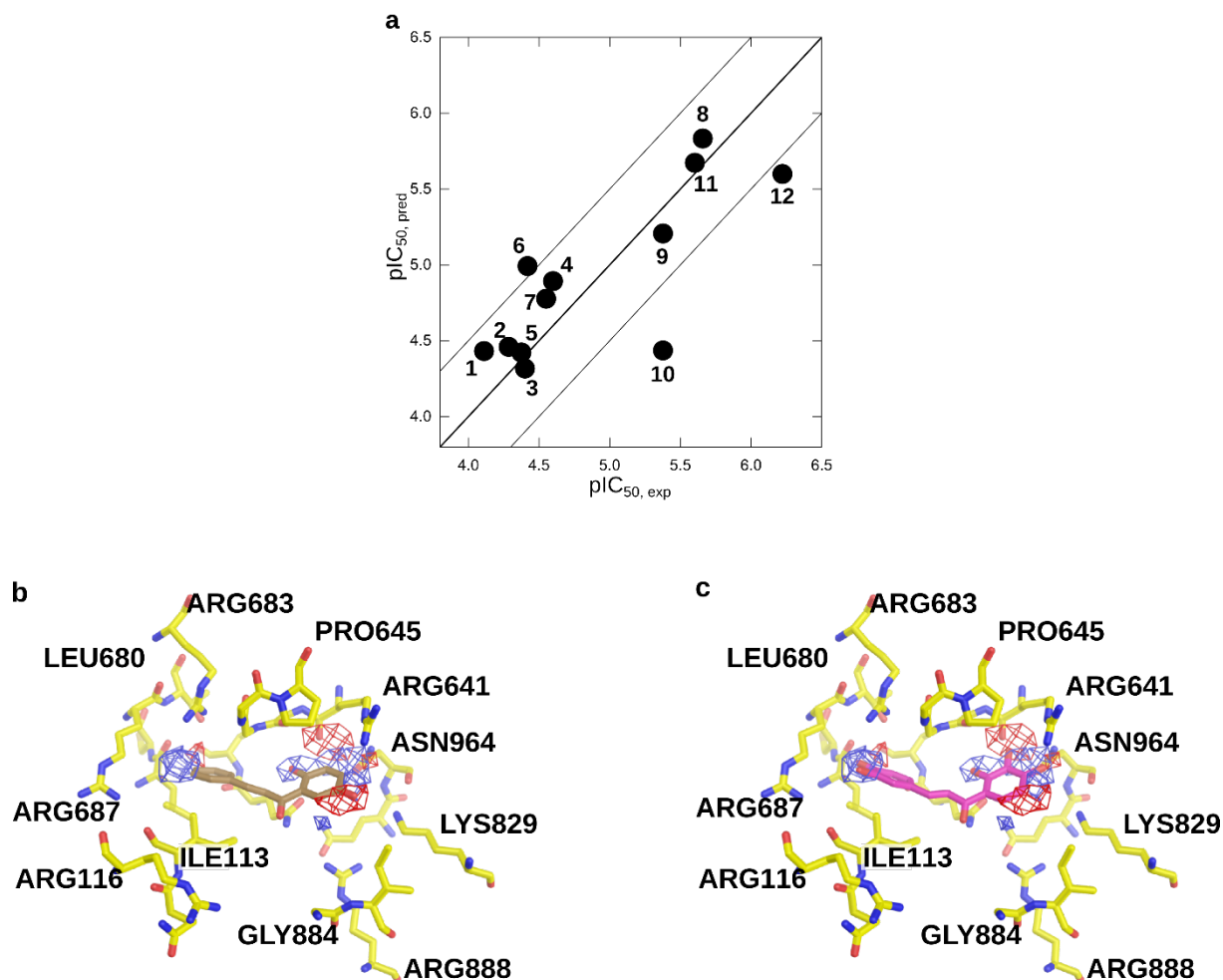
**a**, Overlay of C<sub>4</sub> (yellow) and C<sub>3</sub> (green) PEPC. Residues of the feedback inhibitor binding site are depicted as spheres. **b**, **c**, Redocking of aspartate in C<sub>4</sub> (**b**) and C<sub>3</sub> (**c**) PEPC. The co-crystallized aspartate is shown in grey, and the lowest energy docking pose is shown in orange.



**Figure S3: Common binding motif of tested chalcones and hydrogen bond network.**

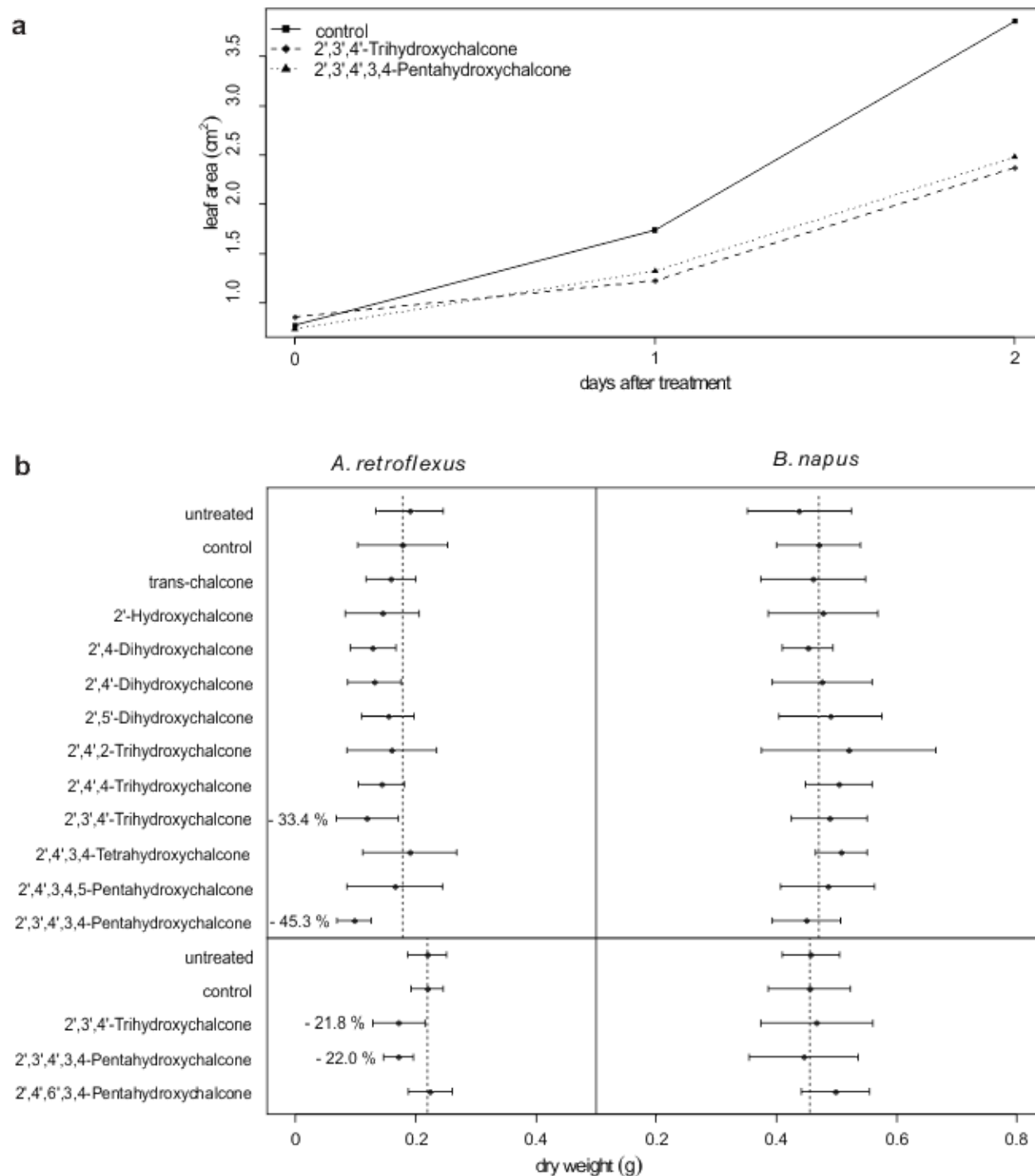
Overlay of binding poses of **8** (grey), **9** (orange), **10** (cyan), **11** (blue), and **12** (magenta) obtained by docking with GLIDE into C<sub>4</sub> PEPC. Dashed lines depict hydrogen bonds and cation- $\pi$ -interactions between the ligands and the protein. The roman numbers label the hydrogen bonds as mentioned in Figure 1d and e.





**Figure S4: Results of the AFMoC analyses for chalcone ligands in  $C_4$ .**

**a**,  $pIC_{50}$  values predicted by leave-one-out analysis versus experimentally determined ones for 12 chalcone ligands (see Table 1 in the main text). In addition to the line of ideal correlation, thin lines are given to indicate deviations of half a logarithmic unit from ideal prediction [24]. **b**, **c**, AFMoC STDEV\*COEFF contour plots elucidating regions in the binding pocket of PEPC where the presence of hydroxyl oxygens will enhance (blue mesh) or reduce (red mesh) binding. Contour levels are -0.0075 and 0.02, respectively. Ligands 5 (**b**, brown;  $pIC_{50} = 4.357$  in  $C_4$  PEPC) and 12 (**c**, magenta;  $pIC_{50} = 6.22$  in  $C_4$  PEPC) are shown for comparison.

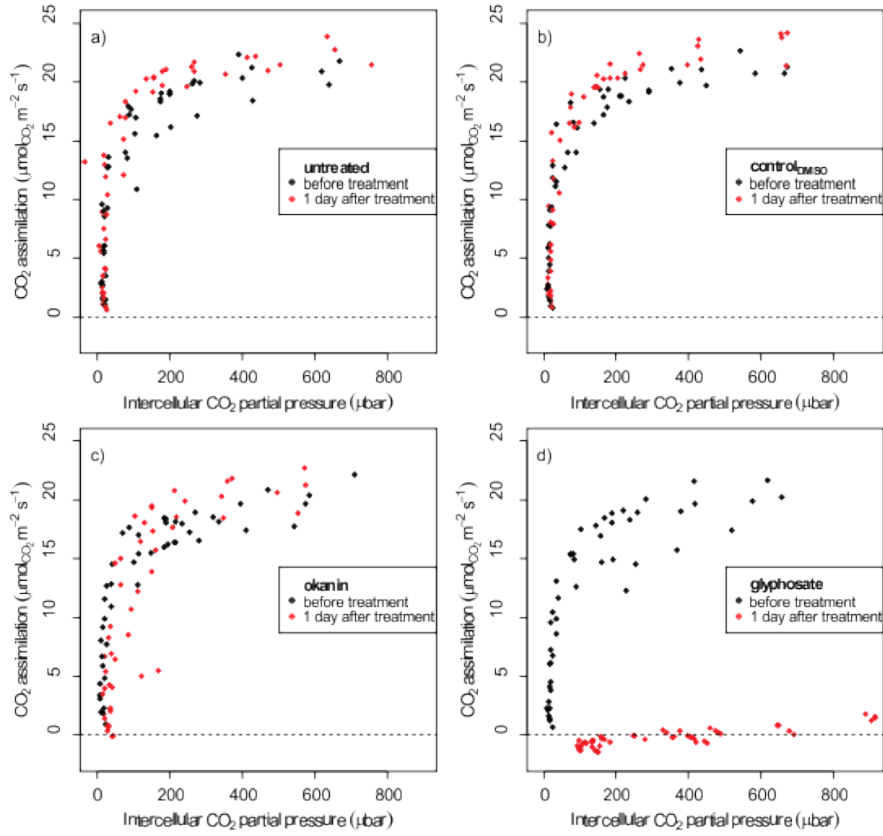


**Figure S5: Influence of selected chalcones on leaf growth of *A. retroflexus* and biomass formation of *A. retroflexus* and *B. napus***

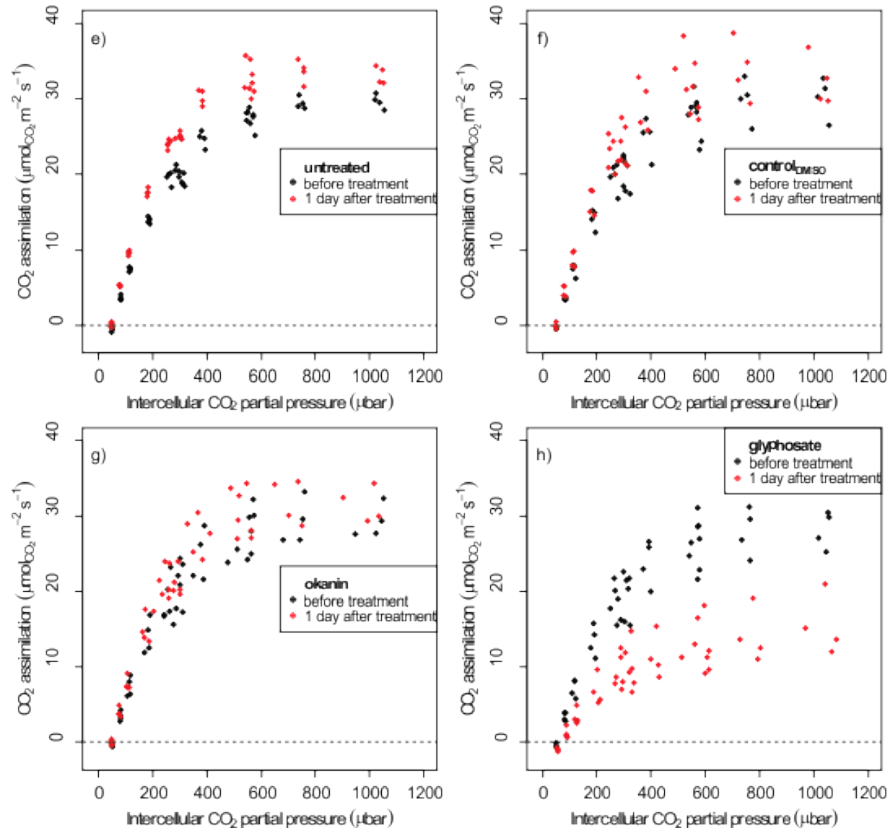
**a**, Growth of the 2<sup>nd</sup> leaf of *A. retroflexus* displayed as leaf area (cm<sup>2</sup>) at 0, 1 and 2 days after treatment with 3 mM 2',3',4'-Trihydroxychalcone (**10**) and 2',3',4',3,4-Pentahydroxychalcone (**12**) in comparison to the control treatment. Rate of growth was significantly reduced for the two displayed chalcones. **b**, Dry weight of *A. retroflexus* and *B. napus* six days after treatment.

Results of experiment 1 (top) and experiment 2 (bottom) are displayed. 2',3',4'-Trihydroxychalcone (**10**) and 2',3',4',3,4-Pentahydroxychalcone (**12**) significantly reduced *A. retroflexus* dry weight, and the effect was not significantly different between the two experiments. None of the compounds had an effect on dry weight of *B. napus*. Bars display standard deviations (n=6; *B. napus* experiment 2: n=5). Dashed lines represent the average of the corresponding control treatments.

## *Amaranthus retroflexus*

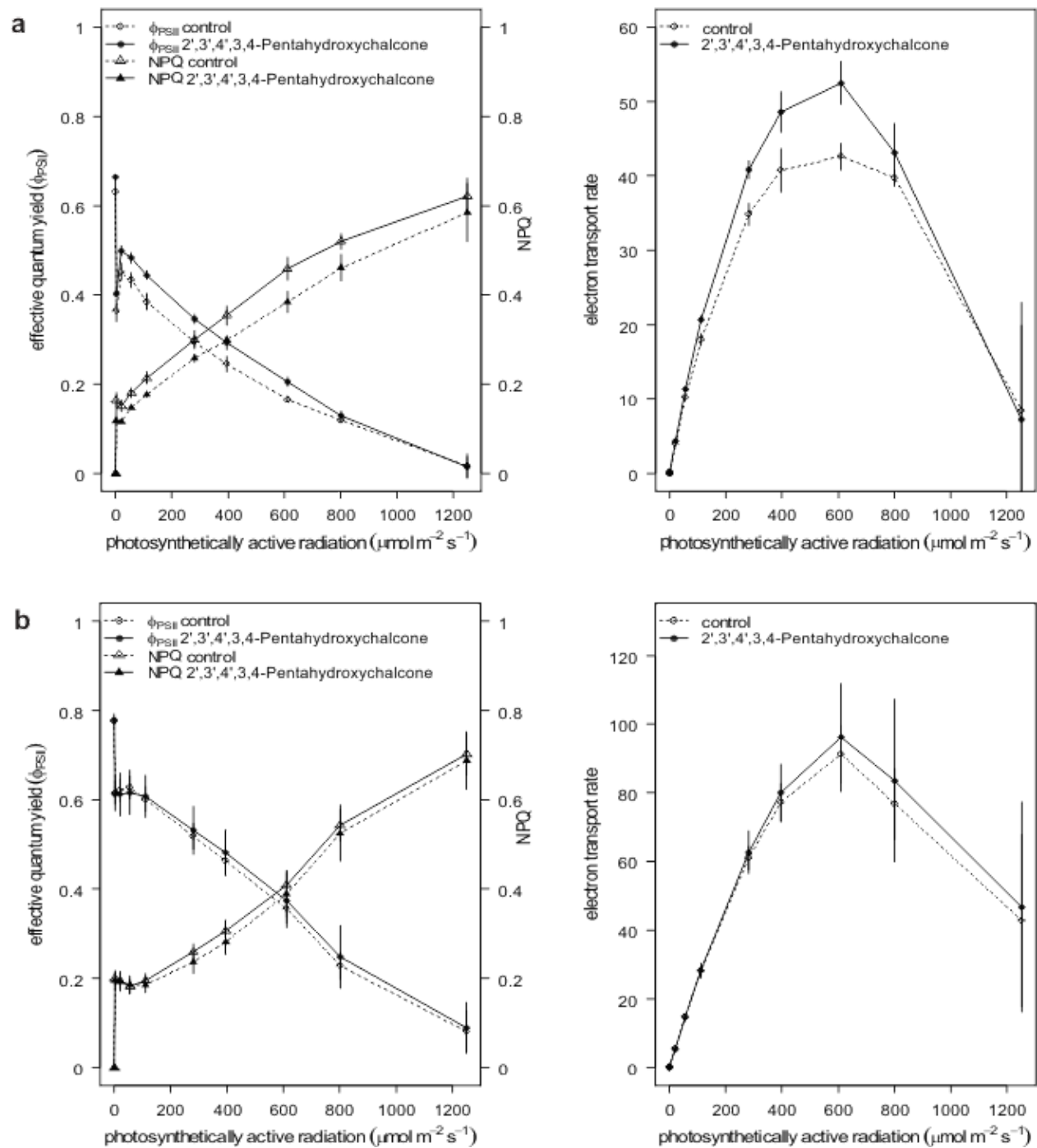


## *Brassica napus*



**Figure S6: Influence of okanin on CO<sub>2</sub> assimilation of *A. retroflexus* (a-d) and *B. napus* (e-h) before treatment and one day after treatment.**

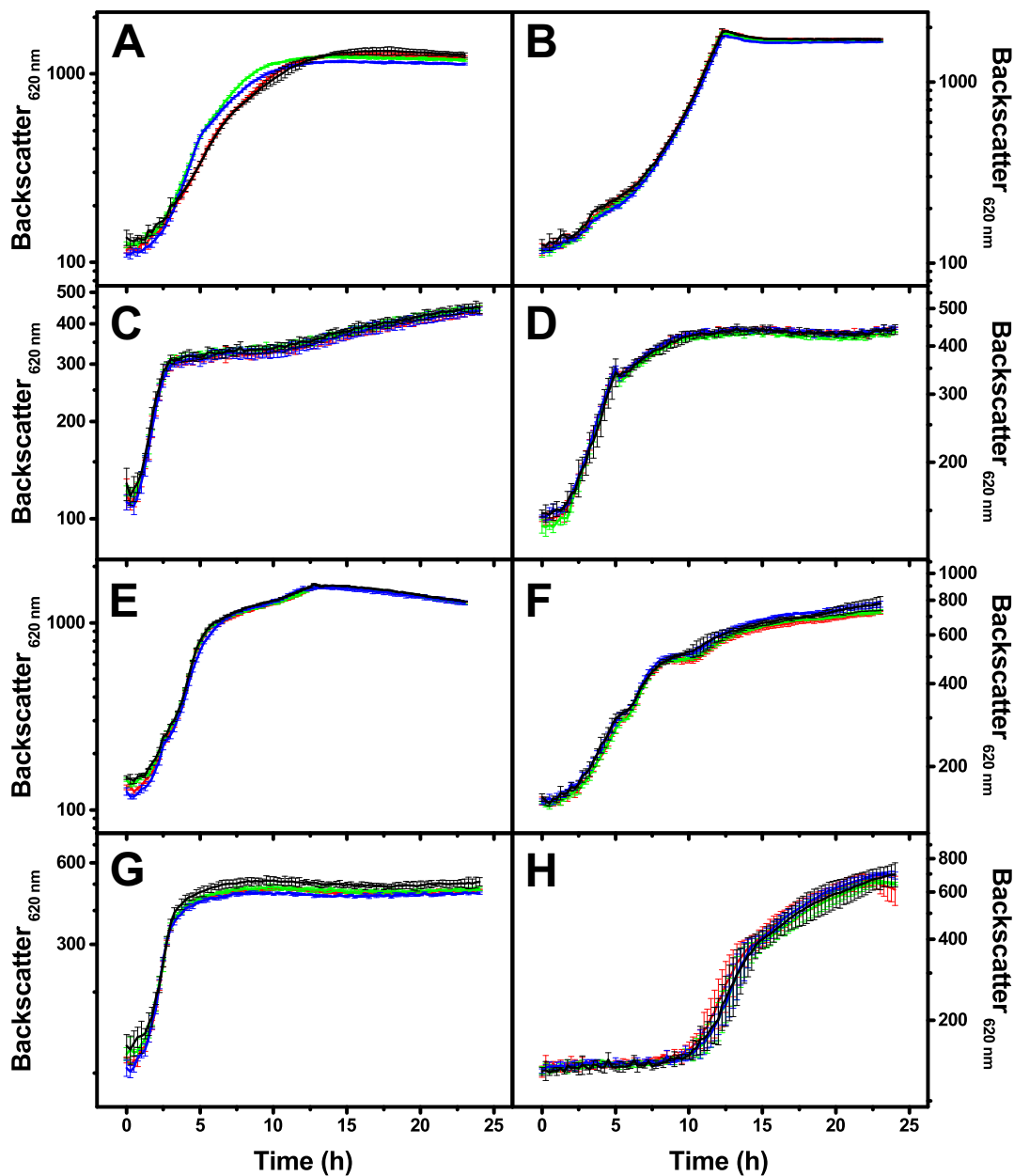
**a, e**, CO<sub>2</sub> assimilation of untreated plants. **b, f**, CO<sub>2</sub> assimilation of control (DMSO) plants. **c, g**, CO<sub>2</sub> assimilation of plants treated with okanin. **d, h**, CO<sub>2</sub> assimilation of plants treated with glyphosate.



**Figure S7: Influence of selected chalcones on photosynthetic parameters of *B. napus* and *A. retroflexus***

**a**, Influence of 2 mM 2',3',4',3,4-Pentahydroxychalcone on photosynthetic parameters of the 2<sup>nd</sup> leaf of *A. retroflexus* measured as rapid light curve in leaf disc assays. Effective quantum yield of photosystem II ( $\Phi_{PSII}$ ) was increased in comparison to the control treatment, whereas PSII non-photochemical quenching (NPQ) was reduced (left). Electron transport rate was increased

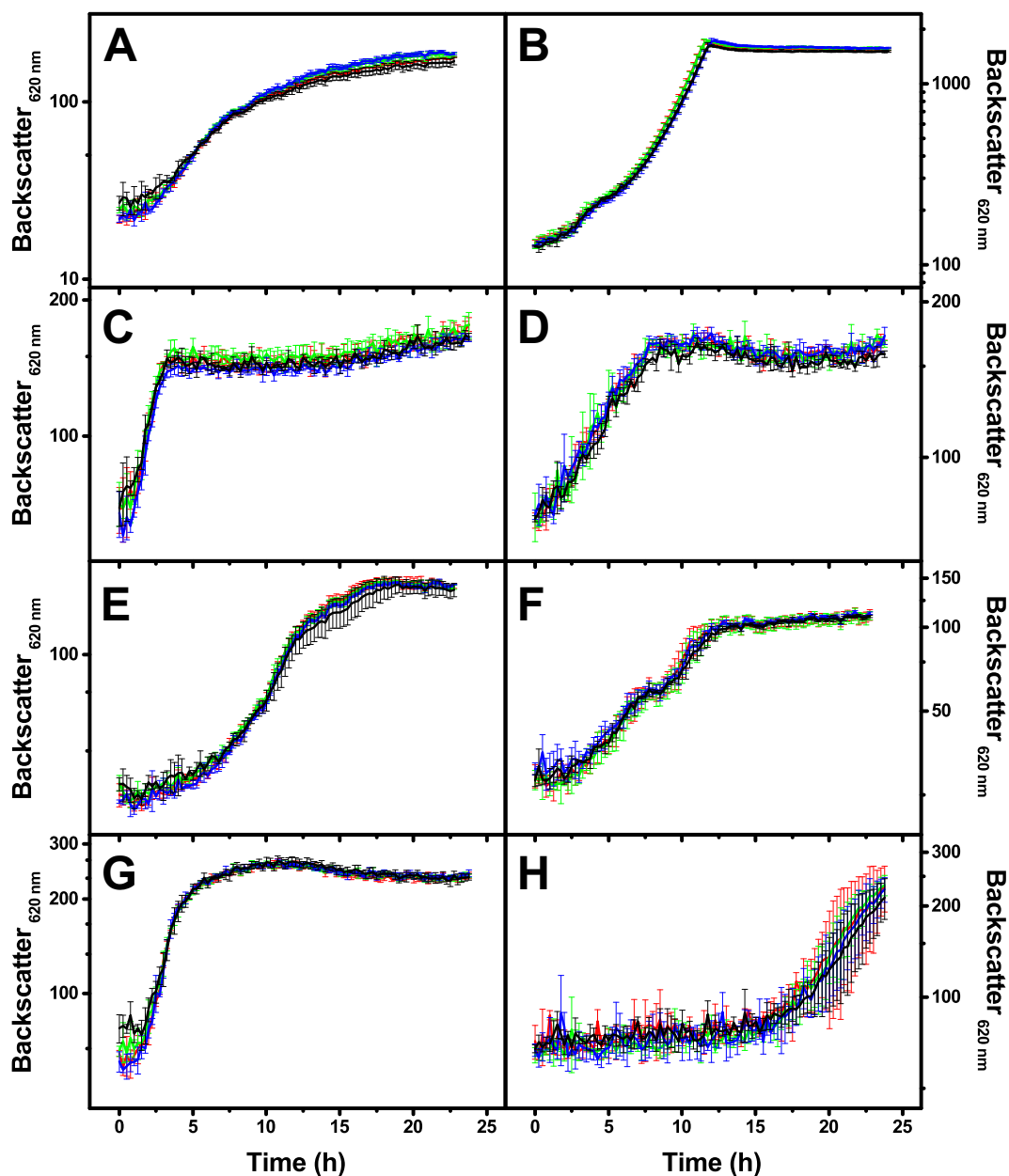
following treatment with 2 mM 2',3',4',3,4-Pentahydroxychalcone. Bars display standard deviation (n=4). **b**, Influence of 2mM 2',3',4',3,4-Pentahydroxychalcone on photosynthetic parameters of the 2<sup>nd</sup> leaf of *A. retroflexus* measured as rapid light curve in leaf disc assays. There was no effect on photosynthetic parameters of *B. napus*. Bars display standard deviation (n=4).



**Figure S8: Cultivation of bacteria in the presence of okanin.**

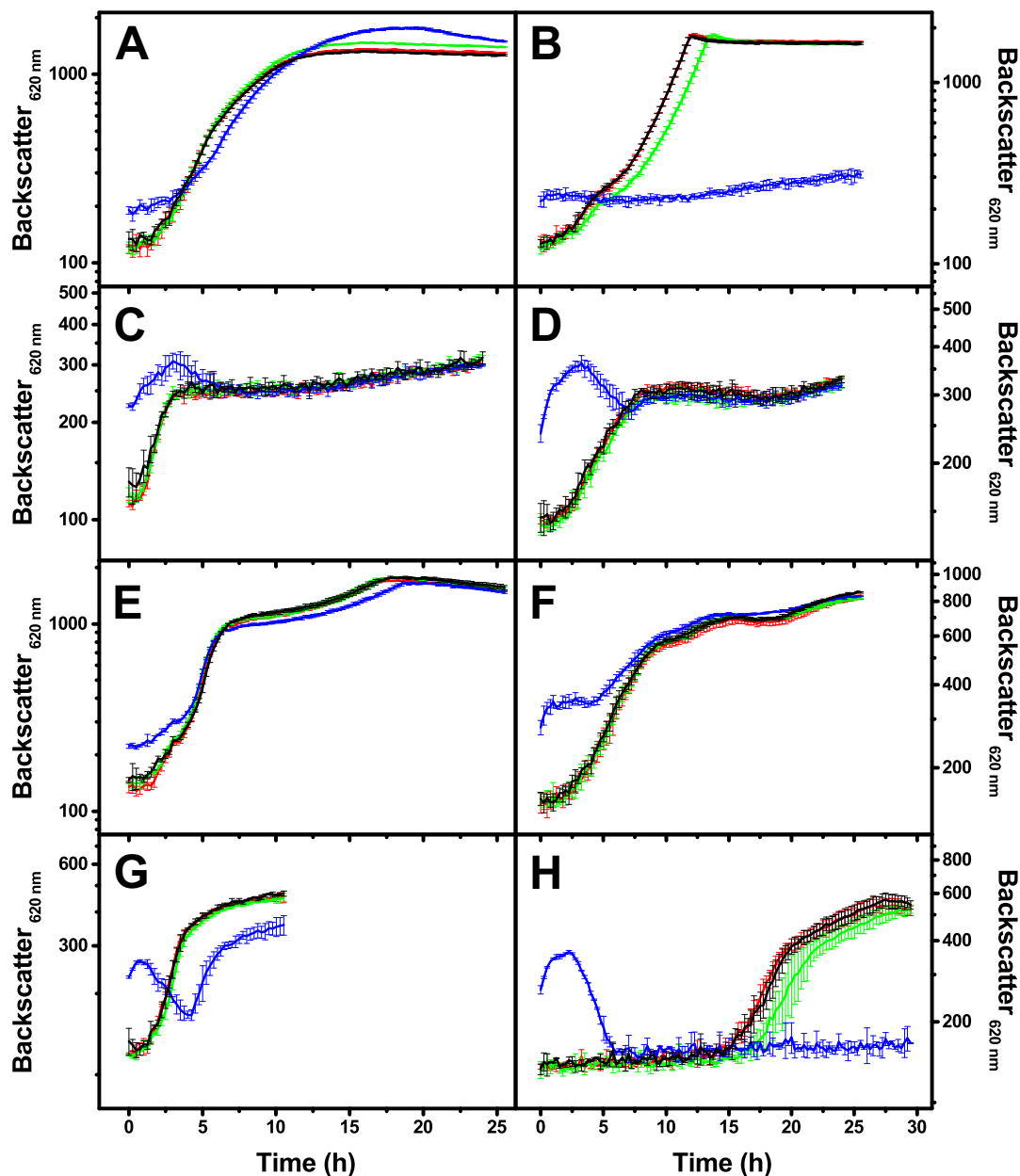
Cultivation of *C. glutamicum* ATCC 13032 (A, B), *E. coli* K-12 MG1655 (C, D), *P. putida* KT2440 (E, F), and *B. subtilis* 168 (G, H) in the presence of 0  $\mu\text{M}$  (black), 0.1  $\mu\text{M}$  (red), 1  $\mu\text{M}$  (green), and 10  $\mu\text{M}$  (blue) okanin (2',3',4',3,4-Pentahydroxychalcone). The cells were cultivated in complex media (BHI medium; A, C, E, G) or minimal media (CGXII, B; M9, D, F; Spizizen, H) containing 2% (*w/v*) glucose. The cell density was monitored in the Biolector system measuring the backscatter at 620 nm (gain 20).





**Figure S9: Cultivation of bacteria in the presence of chalcone 10.**

Cultivation of *C. glutamicum* ATCC 13032 (A, B), *E. coli* K-12 MG1655 (C, D), *P. putida* KT2440 (E, F), and *B. subtilis* 168 (G, H) in the presence of 0  $\mu\text{M}$  (black), 0.3  $\mu\text{M}$  (red), 3  $\mu\text{M}$  (green), and 30  $\mu\text{M}$  (blue) of chalcone **10** (2',3',4'-Trihydroxychalcone). The cells were cultivated in complex media (BHI medium; A, C, E, G) or minimal media (CGXII, B; M9, D, F; Spizizen, H) containing 2% (*w/v*) glucose. The cell density was monitored in the Biolector system measuring the backscatter at 620 nm (gain 20).



**Figure S10: Cultivation of bacteria in the presence of chalcone 5.**

Cultivation of *C. glutamicum* ATCC 13032 (A, B), *E. coli* K-12 MG1655 (C, D), *P. putida* KT2440 (E, F), and *B. subtilis* 168 (G, H) in the presence of 0  $\mu\text{M}$  (black), 5.5  $\mu\text{M}$  (red), 55  $\mu\text{M}$  (green), and 550  $\mu\text{M}$  (blue) chalcone **5** (2',5'-Dihydroxychalcone). The cells were cultivated in complex media (BHI medium; A, C, E, G) or minimal media (CGXII, B; M9, D, F; Spizizen, H) containing 2% (w/v) glucose. The cell density was monitored in the Biolector system measuring the backscatter at 620 nm (gain 20).

## Supplemental Tables

**Table S1:** Parameters of the AFMoC analyses for chalcone ligands in C<sub>4</sub> PEPC.

Spacing <sup>a</sup>	1.0
$\sigma$ <sup>a,b</sup>	0.85
$q^2$ <sup>c,d</sup>	0.63 (0.63)
$s_{press}$ <sup>e,f</sup>	0.10
$r^2$ <sup>c,g</sup>	0.99 (0.99)
$s$ <sup>e,h</sup>	0.08
$F$ <sup>c,i</sup>	175.6 (175.6)
Components <sup>j</sup>	3
Fraction <sup>k,l</sup>	
C.2	0.01
C.ar	0.37
O.3	0.61
O.2	0.01

<sup>a</sup> In Å.

<sup>b</sup> Half-width of the convolution of the receptor potential fields with the Gaussian function representing the ligand atoms.

<sup>c</sup> Values are given considering only the part of pEC<sub>50</sub> used in the PLS analysis (pEC<sub>50</sub><sup>PLS</sup>) or considering the total pEC<sub>50</sub> (values in parentheses).

<sup>d</sup>  $q^2 = 1 - \text{PRESS}/\text{SSD}$  as obtained by “leave-one-out” cross-validation. PRESS equals the sum of squared differences between predicted and experimentally determined binding affinities, SSD is the sum of the squared differences between experimentally determined binding affinities and the mean of the training set binding affinities.

<sup>e</sup> In logarithmic units.

<sup>f</sup>  $s_{press} = \sqrt{\left(\frac{\text{PRESS}}{n-h-1}\right)}$  as obtained by “leave-one-out” cross-validation.  $n$  equals the number of data points,  $h$  is the number of components.

<sup>g</sup> Correlation coefficient.

<sup>h</sup>  $s = \sqrt{\left(\frac{\text{RSS}}{n-h-1}\right)}$ . RSS equals the sum of squared differences between fitted and experimentally determined binding affinities.

<sup>i</sup> Fisher’s F-value.

<sup>j</sup> Number of components.

<sup>k</sup> Fraction of the explained variance due to an interaction field of the given Sybyl atom type.

**Table S2:** Complete growth data of cultivations of *C. glutamicum* ATCC 13032, *E. coli* K-12 MG1655, *P. putida* KT2440, and *B. subtilis* 168 in the presence of chalcone derivatives.

The bacterial strains were cultured in a microscale cultivation system both in complex and minimal media in the presence of 0.1x, 1x, and 10x the IC<sub>50</sub> concentration determined for the *F. trinervia* PEP carboxylase. The table indicates whether the final cell density (measured as backscatter at 620 nm) or the growth rate decreased (fbs↓ and μ↓, respectively), or the lag phase was extended (lag↑). All values are given relatively to an untreated control. The colored background of values indicates a deviation of the mean from the control higher than 10% (excluding standard deviation). Effects due to precipitation in the growth media occurring at 10x IC<sub>50</sub> of *trans*-chalcone and 2'-Hydroxychalcone are indicated by a red background. Abbreviations: StDev, standard deviation; n.d., not determinable.

**Table attached as separate file (Table S2.xlsx)**



**Table S3:** Influence of okanin on maximal rubisco carboxylation rate ( $V_{\text{cmax}}$ ), maximal electron transport rate ( $J_{\text{max}}$ ) and mitochondrial respiration ( $R_d$ ) of *B. napus* estimated from *in planta* gas exchange measurements. Parameter estimates of the Farquhar-vanCaemmerer-Berry-model of the minimum reduced model are given. P-values indicate significance of the F-test comparing reduced and full models. Values in red are final parameter estimates that are valid for both measurement days due to a lack of statistical significance compared with the full model (n=4).

	parameter	before treatment	one day after treatment	<i>p</i>
untreated	$V_{\text{cmax}}$	74.4 (1.82)	94.5 (2.04)	0.001
	$J_{\text{max}}$	136.6 (2.23)	159.0 (2.20)	<0.001
	$R_d$	0.7 (0.35)		1
control <sub>DMSO</sub>	$V_{\text{cmax}}$	162.8 (26.83)	90.9 (4.82)	<0.001
	$J_{\text{max}}$	158.7 (4.79)		0.453
	$R_d$	6.3 (0.93)	1.0 (0.90)	<0.001
okanin	$V_{\text{cmax}}$	81.5 (3.33)		0.2998
	$J_{\text{max}}$	144.2 (3.77)		
	$R_d$	0.9 (0.66)		