

**Crystalline Sponges as a sensitive and fast method for metabolite identification:**

**Application to gemfibrozil and its phase I and II metabolites**

Lara Rosenberger, Carolina von Essen, Anupam Khutia, Clemens Kühn, Klaus Urbahns,  
Katrini Georgi, Rolf W. Hartmann, Lassina Badolo

Department of Drug Metabolism and Pharmacokinetics (DMPK), Merck KGaA, Frankfurter  
Strasse 250, 64293 Darmstadt, Germany (L.R., K.U., K.G., L.B.)

Innovation Center, Merck KGaA, Frankfurter Strasse 250, 64293 Darmstadt, Germany (C.v.E.,  
A.K., C.K.)

Department of Drug Design and Optimization (DDOP), Helmholtz-Institute for Pharmaceutical  
Research Saarland (HIPS)-Helmholtz Centre for Infection Research (HZI), Campus E8.1,  
66123 Saarbrücken, Germany (L.R., R.W.H.)

**Running Title:** Complete structural elucidation of gemfibrozil metabolites by crystalline sponges

**Address correspondence to:**

Dr. Lassina Badolo,

Merck KGaA,

Frankfurter Str. 250, 64293 Darmstadt, Germany

E-mail: lassina.badolo@emdserono.com

Number of Text Pages: 20

Number of Tables: 0

Number of Figures: 6

Number of References: 28

Number of words in Abstract: 236

Number of words in Introduction: 753

Number of words in Discussion: 584

**Abbreviations:** API, active pharmaceutical ingredient; CS, crystalline sponge; CS-XRD, crystalline sponge method; DME, 1,2-dimethoxyethane; HPLC, high performance liquid chromatography; LC, liquid chromatography ; MS<sup>2</sup>, tandem mass spectrometry; MS, mass spectrometry; SC-XRD, single crystal X-ray diffraction; tpt, 2,4,6-tris(4-pyridyl)-1,3,5-triazine; UDPGA, uridine 5'-diphosphoglucuronic acid; UPLC, ultra performance liquid chromatography

## Abstract

Understanding the metabolism of new drug candidates is important during drug discovery and development, as circulating metabolites may contribute to efficacy or cause safety issues. In early phase of drug discovery, human in vitro systems are used to investigate human relevant metabolism. While conventional techniques are limited in their ability to provide complete molecular structures of metabolites (liquid chromatography mass spectrometry (LC-MS)) or require larger amount of material not available from in vitro incubation (nuclear magnetic resonance (NMR)), we here report for the first time the use of the crystalline sponge method (CS-XRD) to identify phase I and phase II metabolites generated from in vitro liver microsomes or S9 fractions. Gemfibrozil was used as a test compound. Metabolites generated from incubation with microsomes or S9 fractions, were fractionated using online fraction collection. After chromatographic purification and fractionation of the generated metabolites, single crystal X-ray diffraction (SC-XRD) of crystalline sponges (CS) was used to identify the structure of gemfibrozil metabolites. This technique allowed for complete structure elucidation of 5'-CH<sub>2</sub>OH gemfibrozil (M1), 4'-OH gemfibrozil (M2), 5'-COOH gemfibrozil (M3) and the acyl glucuronide of gemfibrozil, 1-O-β-glucuronide (M4), the first acyl glucuronide available in the Cambridge Crystallographic Data Centre. Our study shows that when optimal soaking is possible, crystalline sponges technology, is a sensitive (ng amount) and fast (few days) method that can be applied early in drug discovery to identify the structure of pure metabolites from in vitro incubations.

### **Significance Statement**

Complete structure elucidation of human metabolites plays a critical role in early drug discovery. Low amounts of material (ng) are only available at this stage and insufficient for NMR analysis. The crystalline sponge method has the potential to close this gap, as demonstrated in this study.

## Introduction

During the process of drug discovery and development, assessing the identity of drug metabolites is crucial to understand the contribution of metabolites to efficacy or safety (Nedderman, 2009; Zhang et al., 2009). Guidelines have been drafted by regulatory agencies (EMA, 2013; FDA, 2016) to help the pharmaceutical industry to apply the right strategy regarding metabolites in safety assessment. The key element in all these guidances is to protect patients and healthy human volunteers from potentially toxic metabolites with a focus on human metabolites that are either absent in preclinical species or circulate in these animals at much lower concentrations than in humans (Ma and Chowdhury, 2011). It is therefore of paramount importance during drug development to gain an understanding of metabolism of new chemical entities in human as early as possible. To achieve this, in vitro systems from human origin (e.g., microsomes or hepatocytes) are used to metabolize the new active pharmaceutical ingredients (API) in drug discovery research (Brandon et al., 2003; Fasinu et al., 2012). Samples generated from these systems are analyzed with LC-MS of different types (e.g., tandem mass spectrometry (MS<sup>2</sup>) and tandem-time-of-flight) by drawing inferences from fragmentation patterns to assess metabolism (Prakash et al., 2007; Prasad et al., 2011). Although MS techniques have reached extreme sensitivity, these techniques fail in many situations, e.g., to identify the complete structure of metabolites. This forces scientists to use NMR, a less sensitive technique, requiring larger amounts of material (usually µg to mg range) (Murai et al., 2004), which is therefore not always suitable for analyzing samples from in vitro origin.

Besides NMR and LC-MS, another technique would be single crystal X-ray diffraction (SC-XRD), a direct technique that provides structural information at the atomic level by measuring electron density as diffraction pattern of a single crystal (Massa, 2004). SC-XRD can even provide information on the absolute configuration of chiral centers but has the intrinsic limitation due to requirements of single crystals of suitable sizes and shape. The process of crystallization is a time-consuming procedure, that requires availability of sufficient amounts of materials and cannot be performed with amorphous solids, liquids or volatile analytes. One

approach to overcome the limitation of crystallization is the application of microcrystal electron diffraction from powder (Jones et al., 2018; Kunde and Schmidt, 2019). This technique requires solid material at higher amount.

For structural elucidation of metabolites in solution, a method with the information density of X-ray crystallography and the sensitivity in the ng range is needed for human related metabolites at this early stage of drug development research. In 2013, Makoto Fujita published a new technique commonly known as “crystalline sponge method”, which enables crystal structure determination without crystallization with only ng to few µg amount of analytes (Inokuma et al., 2013). In one specific case only 80 ng of material were sufficient to clearly elucidate the X-ray structure of the analyte and presumably even smaller amounts are possible for synchrotron X-ray diffraction instruments. The CS-XRD uses pre-existing crystals of porous metal coordination complexes  $[(ZnX_2)_3 \cdot (tpt)_2 \cdot x(\text{solvent})]_n$  ( $X = Cl$ ;  $tpt = 2,4,6$ -tris(4-pyridyl)-1,3,5-triazine), which can absorb organic molecules in its pores and make them observable by conventional single crystal X-ray crystallography. Via diffusion (“soaking”) the analyte (“guest”) is absorbed into the pores of porous CS and regularly ordered by intermolecular, non-covalent interactions (“guest soaking”) (Inokuma et al., 2013; Inokuma et al., 2016; Sakurai et al., 2017). As a result, the repetitive positioning of the analytes in each pore of the framework serves for structural analysis via X-ray diffraction. For a successful CS-XRD experiment regular order of the analyte is important. Functionalization of the API’s (e.g., hydroxylation, demethylation) increases the type of non-covalent interactions with the framework and therefore increases the chances of regular order. This even works for otherwise non-crystallizable compounds (Hoshino et al., 2016). The CS-XRD has been shown to be able to successfully elucidate the structure of API’s (Sakurai et al., 2017), products from biosynthesis catalyzed by cytochrome P450 enzymes (Morita et al., 2019), natural plant products (Wada et al., 2018) and metabolic products prepared by reductive dichlorination of an insecticide as well as keton reduction of a steroid hormone with baker’s yeast (Inokuma et al., 2016), but has not been applied for drug-related material generated by incubation of API’s with human or animal tissues before.

In this study, we apply the crystalline sponge technology for complete structure identification of gemfibrozil phase I and phase II metabolites after incubation with liver microsomes or S9. Rat systems were used, however, the formation of the same metabolites were confirmed in human liver S9.

## Materials and Methods

**Chemicals and Reagents.** Gemfibrozil (Figure 1) was purchased from Acros Organics (New Jersey, USA) and its metabolite 4'OH gemfibrozil was synthesized from WuXi AppTec (Wuhan) Co., Ltd. (Wuhan, China). Gemfibrozil glucuronide was purchased from Toronto Research Chemicals (Toronto, Canada). Di-potassium hydrogen phosphate, potassium dihydrogen phosphate, magnesium chloride hexahydrate, dimethylsulfoxide, cyclohexane, methanol, nitrobenzene, nicotinamide adenine dinucleotide phosphate disodium salt (NADP- $\text{Na}_2$ ), water (UHPLC-MS grade) and acetonitrile (UHPLC-MS grade) were purchased from Merck KGaA (Darmstadt, Germany). Zinc chloride, 1,2-dimethoxyethane (DME), alamethicin, uridine 5'-diphosphoglucuronic acid trisodium salt (UDPGA- $\text{Na}_3$ ) and formic acid were purchased from Sigma Aldrich Chemie GmbH (Steinheim, Germany). Dihydronicotinamide adenine dinucleotide phosphate tetrasodium salt (NADPH- $\text{Na}_4$ ) was purchased from AppliChem (Darmstadt, Germany) and 2,4,6-tri(4-pyridyl)-1,3,5-triazine (tpt) was purchased from abcr GmbH (Karlsruhe, Germany). Female WistarHan rat liver microsomes (R6500, pool of 225), female rat liver S9 (R3500.S9, pool of 100) and mixed gender human liver S9 (H0620.S9, pool of 50) were obtained from Sekisui XenoTech (Kansas City, USA) and mixed gender human liver microsomes (Ultrapool, pool of 150) were purchased from Corning (Corning, USA)

**Metabolism of gemfibrozil by rat/human liver microsomes, rat/human liver S9.** The oxidation reaction of gemfibrozil was conducted with rat liver microsomes (0.5 mg/mL) in 50 mM potassium phosphate buffer (pH 7.4) containing 1 mM magnesiumchloride and 20  $\mu\text{M}$ /200  $\mu\text{M}$  substrate. After 5 min of preincubation (37 °C, 150 rpm), the reaction was initiated by the addition of NADPH (1.5 mM) dissolved in 50 mM potassium buffer and the mixture was incubated for another 3 hours (final volume: 5.40 mL). The reaction was terminated by adding one volume of cold acetonitrile and then centrifuged (4.000 g, 1 h, 4 °C) to sediment the precipitated proteins. The aliquots (100  $\mu\text{L}$ ) of the supernatant were injected onto high performance liquid chromatography coupled with mass spectrometry (HPLC-MS) for separation and fractionation.



The oxidation of gemfibrozil with rat and human liver S9 fraction was conducted similar to the incubation with microsomes. Differences were the protein concentration of 1.0 mg/mL, the addition of NADPH and NADP<sup>+</sup> (1.5 mM) as cofactors and the incubation time (6 h).

To generate gemfibrozil glucuronide, human liver microsomes were treated with alamethicin (4 °C, 20 min, final concentration: 25 µg/mL), a channel forming peptide to increase the metabolite formation. The reaction was initiated by the addition of UDPGA (1.5 mM) as cofactor.

**HPLC-MS.** The supernatants were analyzed on an Acquity Arc HPLC system combined with an isocratic solvent manager, a single quadrupole MS and a fraction manager (Waters Corporation, Milford, USA). The LC system included a quaternary solvent manager, a sample manager and a 2998 photodiode array detector. The isocratic solvent manager was used with a 1:50 splitter. Samples were analyzed with electrospray ionization mass spectrometry in the negative ion mode. The following source parameters were used: probe temperature 600 °C; capillary voltage 0.8 V; cone voltage 3.0 V and sampling frequency 2 Hz. Ions were acquired in an MS acquisition range from m/z 100 to m/z 650 in continuum mode.

HPLC separations were achieved on two Chromolith Performance RP18e columns as first chromatographic system (100-4.6 mm; Merck KGaA, Darmstadt, Germany) and Purospher Star RP18e Hibar HR column as second system (100-2.1 mm, 2 µm; Merck KGaA, Darmstadt, Germany) at a column oven temperature of 25 °C. For the first purification elution was performed at a flow rate of 1.0 mL/min over a period of 30 minutes with a mixture of solvent A (water + 0.1% formic acid) and solvent B (acetonitrile + 0.1% formic acid). Compounds were eluted using the following conditions: mobile phase B increased from 0% to 25% over 7 min, followed by a 13 min linear gradient to 60% B and a linear gradient to 100% B in 5 min; returning to 0% B in 0.1 minutes and re-equilibration at 0% B for 4.9 min. The first 6 min of eluent were directed to waste to reduce contamination, as it contains mainly microsomes and buffers salts from incubation matrix. For the second purification elution was performed at a flow rate of 0.45 mL/min over a period of 15 minutes using the following conditions: mobile

phase B increased from 0% to 80% over 10 min, followed by linear gradient to 100% B in 2 min; returning to 0% B in 0.1 minutes and re-equilibration at 0% B for 2.9 min. The eluate was subsequently split between the photodiode array detector and fraction manager (98%), and the mass spectrometer (2%). The MS flow was increased to 0.3 ml/min with a makeup solvent (90% water + 10% acetonitrile). The fractionation of the analytes was triggered by the mass values ( $m/z$  265,  $m/z$  279 and  $m/z$  425) registered by the MS detector. Data acquisition and sample fractionation were performed using the software MassLynx 4.2 combined with FractionLynx.

Collected fractions were pooled, evaporated to dryness under nitrogen flow at 40°C and resolubilized in acetonitrile/water (40%/60%) for further purification. The fractions from the second purification step were evaporated to dryness for the soaking process.

**UPLC-MS<sup>2</sup>.** The supernatants were analyzed on an Acquity UPLC Class I system (Waters Corporation, Milford, USA) combined with a linear ion trap quadrupole (Qtrap 5500) mass spectrometer (AB Sciex LLC, Framingham, USA). Samples were analyzed with an electrospray ion source in the negative ion mode. Source parameters were source temperature 600°C and ion spray voltage -4.5 kV. UPLC separations were achieved on an Acquity UPLC BEH C18 column (1.7  $\mu$ m; 2.1 x 50 mm; Waters Corporation, Milford, USA) at a column oven temperature of 40 °C. Elution was performed at a flow rate of 0.7 ml/min over a period of 5.2 minutes with a mixture of solvent A (water + 0.1% formic acid + 10 mM ammonium formate) and solvent B (acetonitrile). Compounds were eluted using the following conditions: mobile phase B increased from 30% to 50% between 0.1 min and 4.0 min, followed by a 0.6 min linear gradient to 100% B, returning to 30% B in 0.02 minutes and re-equilibration at 30% B for 0.58 min. After injection, the column effluent was directed to the waste for 0.2 min, to reduce contamination. The software Analyst 1.6.3 was used for data acquisition.

**Crystalline sponge method.** The porous crystalline sponge  $[(\text{ZnCl}_2)_3(\text{tpt})_2 \cdot (\text{cyclohexane})_x]_n$  was prepared according to the reported procedure (Biradha et al., 2002; Ramadhar et al., 2015).

For guest soaking one of the prepared  $[(\text{ZnCl}_2)_3(\text{tpt})_2 \cdot x(\text{cyclohexane})]_n$  single crystals was used as crystalline sponge, and then transferred with 50  $\mu\text{L}$  cyclohexane to the sample vial. Reference material was prepared as 1 mg/mL solution in DME, 1  $\mu\text{L}$  of the solution pipetted into the sample vial, the vial closed with a screw cap with septum seal (VWR, Darmstadt, Germany) and pierced with a syringe needle for slow evaporation of the solvent during guest inclusion. The samples were placed in an incubator at 50 °C for 1 day. Soaking conditions for gemfibrozil glucuronide were conducted similarly. The volume of cyclohexane was reduced to 20  $\mu\text{L}$  and the incubator temperature was 25 °C. Incubation samples were prepared similarly to the reference material. Differences were that 40  $\mu\text{L}$  cyclohexane was transferred together with the crystal into the glass vial, which contained the pooled metabolite material. 4  $\mu\text{L}$  of DME was added subsequently. The guest soaking for gemfibrozil glucuronide was performed using 20  $\mu\text{L}$  of cyclohexane at 25 °C for 1 day.

Single crystal X-ray diffraction measurements were conducted using a Rigaku Oxford Diffraction XtaLAB Synergy-R diffractometer using Cu-K $\alpha$  X-ray radiation ( $\lambda = 1.54184 \text{ \AA}$ ), equipped with a HyPix-6000HE/HyPix-Arc 150° Hybrid Photon Counting (HPC) detector (Rigaku, Tokyo, Japan) at a temperature of 100 K using a Cryostream 800 nitrogen stream (Oxford Cryostreams, UK). The software CrysAlisPro (Rigaku Oxford Diffraction, 2018) was used for calculation of measurement strategy and data reduction (data integration, empirical and numerical absorption corrections and scaling). Crystal structures were modeled using OLEX2 (Dolomanov et al., 2009), solved with SHELXT and refined using SHELXL (Sheldrick, 2015). First, the refinement of the framework was performed by assigning atoms, modeling of disorder, applying anisotropic refinement and adding of H atoms. The electron density of the guest molecules could safely be identified in the residual electron density after refinement of the framework. The amount of analyte in the sponge pores is modeled as occupancy of the

respective atom coordinates, which is proportional to the observed electron density. The data quality of the measurement was assessed by  $R_{\text{int}}$  and the refined model of framework and analyte by  $R_1$ . The  $R_1$  value represents the agreement between calculated and observed model, and the  $R_{\text{int}}$  is the measure of precision/reproducibility, classifying the recorded reflections from different angles.

The complete process from analyte generation to processing crystallographic data is shown in Figure 2.

## Results

**LC-MS analysis of gemfibrozil metabolites prepared by incubation with rat/human liver microsomes and rat/human liver S9.** Gemfibrozil (Figure 1) was chosen as a tool compound to assess the ability of the CS to elucidate the complete structures of metabolites obtained from in vitro incubation. Hydroxylation reactions were conducted by hydroxylases present in microsomes and further oxidation to carboxylic acid was performed by alcohol and aldehyde dehydrogenases present in the cytosol part of S9 fraction. Formation of gemfibrozil glucuronide was conducted by uridine diphosphoglucuronosyltransferase present in microsomes through conjugation with glucuronic acid (Zhang and Surapaneni, 2012) (Figure 3).

After incubating gemfibrozil, the sample solution was tested with UPLC (ultra performance liquid chromatography)-MS-time-of-flight for metabolite identification. The settings applied on the MS-time-of-flight, did not allow for structure identification of neither the parent nor the metabolite. Further optimization was skipped, and the samples were rather analyzed on single quadrupole HPLC-MS and UPLC-MS<sup>2</sup> in multiple-reaction monitoring mode.

The HPLC-MS chromatogram of gemfibrozil, incubated with rat liver microsomes and liver S9 fraction in the presence of NADPH or NADPH/NADP<sup>+</sup>, showed the formation of two hydroxy- (M1 and M2) and one carboxy gemfibrozil metabolite (M3). The three metabolites could not be observed in the absence of the cofactors, indicating that the metabolites were formed by enzymatic oxidation. All three metabolites detected in rat system were present in human liver S9 (Figure 4). Gemfibrozil incubated with human liver microsomes in the presence of UDPGA, showed the formation of a glucuronide (M4), which was only observable by addition of the cofactor. The metabolites M1 and M2 were detected at a deprotonated molecular mass [M-H]<sup>-</sup> of m/z 265. The m/z value was 16 Da higher than the [M-H]<sup>-</sup> ion of gemfibrozil (m/z 249), suggesting the addition of one oxygen to the molecular structure. M3 was detected as deprotonated molecule [M-H]<sup>-</sup> at m/z 279, suggesting a further addition of oxygen and loss of two hydrogens, but giving no information about the site of metabolism. M4 was detected at a

deprotonated molecular mass  $[M-H]^-$  of  $m/z$  425, indicating the addition of a glucuronic acid (176 Da).

To obtain more information, gemfibrozil and its metabolites were analyzed with UPLC-MS<sup>2</sup>. Tuning of instrument parameters provided the optimal transition for 4'-OH gemfibrozil reference material and was also applied for detection of M1 gemfibrozil. Precursor ion and product ion were adjusted to measure M3 gemfibrozil. Collision induced dissociation of gemfibrozil and its glucuronide generated a main product ion of  $m/z$  121.0 (2',5'-dimethylphenoxy moiety). As a result of the loss of the glucuronide moiety, both compounds were measured with identical transition ( $m/z$  249.1/121.0), but could be differentiated by their retention time. Hydroxy gemfibrozil and carboxy gemfibrozil were detected with  $m/z$  values of 137.2 and 151.0, respectively (Figure 4), indicating that oxidation (+16 Da) and carboxylation (+30 Da) occurred at the 2',5'-dimethylphenoxy moiety and not at the chain of the carboxyalkyl group. This knowledge allowed to limit the metabolic position, but could not give a more precise answer. However, CS-XRD provided precise information about the oxidation and glucuronidation sites as well as the structure of the metabolites.

**Structural elucidation of M1, M2, M3 and M4 gemfibrozil by CS-XRD.** A successful structure determination using the crystalline sponge method depends on various parameters such as temperature, analyte concentration, duration of the soaking experiments and solvent combinations during the soaking step (Hoshino et al., 2016). These parameters need to be optimized for each analyte. Different combinations of these parameters should be tried during this soaking process. Therefore, first analysis was conducted with 1000 ng for each experiment of commercially available gemfibrozil, 4'-hydroxy gemfibrozil and gemfibrozil 1-O- $\beta$ -glucuronide reference material to find the optimal soaking conditions for our tool compounds. The optimized soaking conditions were then applied to its metabolites generated from incubation. The reference material of 4'-hydroxy gemfibrozil was initially used to examine if this technology can be used for polar hydroxy and glucuronide metabolites.

Both, the parent API and the reference metabolites 4'-hydroxy gemfibrozil (Figure 5B) and gemfibrozil 1-O- $\beta$ -glucuronide could successfully be structurally elucidated by CS-XRD using an in-house XRD system. The crystal structure of gemfibrozil revealed two molecules with occupancies of 100% and the final  $R_1$  and  $R_{int}$  values were 11.25% and 1.81%, respectively. The crystal structure of gemfibrozil 1-O- $\beta$ -glucuronide revealed one molecule with occupancy of 100% and final  $R_1$  and  $R_{int}$  values of 6.99% and 1.87%, respectively. Due to disorder, conformation of the glucuronide, i.e. the oxane ring and its substituents cannot be determined. The presence of aliphatic rings in unfavorable conformations is often observed in the CS-XRD, especially for cyclohexane molecules. Nevertheless, the position of glucuronidation at the carboxylic acid function of gemfibrozil can be confirmed.

The successful results demonstrate that the application of the CS method can be expanded to the compounds containing highly polar hydrophilic groups introduced through metabolism of API's, e.g., using cytochrome P450 enzymes, despite the hydrophobic pore of the CS.

Next, the CS method was applied to metabolites obtained by incubation. Incubating gemfibrozil with rat liver microsomes for M1 and M2, rat liver S9 for M3 and human liver microsomes for M4 resulted in the formation of four metabolites in different quantities as described above. The reference material of 4'-OH gemfibrozil and gemfibrozil 1-O- $\beta$ -glucuronide allowed the quantification of M2 and M4, whereas the amounts of M1 and M3 were estimated. For pooling of M1, M3 and M4 an incubation solution of 20  $\mu$ M gemfibrozil was used and 200  $\mu$ M gemfibrozil for M2. The direct use of the separated metabolites was not possible due to the sensitivity of the CS-XRD against impurities as already mentioned in previous publications (Inokuma et al., 2016; Kai et al., 2018). The impurities resulting from the incubation matrix, that is necessary for preparation of metabolites by enzymatic reaction (e.g., salts, cofactor, microsomal stock solution), could not be separated using only one purification step. Therefore, the pooled samples were further purified using a second column. CS soaking experiments were conducted with final amounts in a range of 500 to 1000 ng per experiment.

The soaking experiments of the obtained metabolites M1-M4 were carried out under the same conditions as used for the reference material. The soaked CS were then measured via XRD

and the crystallographic analysis clearly revealed the positions of hydroxylation and carboxylation. M1 exhibits relatively high electron density in close proximity to the 5'-methyl carbon atom forming a benzyl alcohol, which is further oxidized at the same position to form a benzoic acid (M3). M2 shows high electron density at position 4' of the 2',5'-dimethylphenoxy moiety. The electron density map of the benzene core from gemfibrozil and its metabolites M1-M3 are shown in Figure 5B. Two molecules of 5'-hydroxymethyl gemfibrozil and one cyclohexane were clearly assigned by its electron density in the asymmetric unit with occupancies of 100% (Figure 5A). The modeled structure shows a  $R_1$  value of 7.58% and a value of 2.18% for  $R_{int}$ . One guest is stabilized by  $\pi$ - $\pi$  stacking interactions between its aromatic ring and the tpt ligand of the framework [distance 3.63 Å, angle 5.83 °]. Crystallographic data of 4'-hydroxy gemfibrozil revealed one analyte molecule with occupancy of 44% ( $R_1$  7.56%,  $R_{int}$  2.09%) and two cyclohexane molecules. M2 interacts with the framework by  $\pi$ - $\pi$  stacking interactions between the aromatic ring and the tpt ligand [distance 3.54 Å, angle 7.90 °], as well as CH $\cdots$ Cl interactions [distance 2.50 Å] between the 5'-methylgroup of the analyte and the ZnCl<sub>2</sub> part of the framework (Figure 6). The data of 5'-carboxy gemfibrozil provided low electron density of the analyte but clearly confirmed the position of metabolism in comparison with the crystal structure obtained from 5'-hydroxymethyl gemfibrozil (Figure 5B). The XRD data of M4 showed the formation of an acyl glucuronide by conjugation of the carboxylic acid moiety of gemfibrozil with the C1-hydroxy group of glucuronic acid. This confirmed the position of glucuronidation, but due to low electron density of the analyte, the reference material was used to illustrate the structure in this publication. M3 and M4 could be assigned with low electron density compared to the hydroxy metabolites because of the higher polar and nucleophilic groups. The refined structures of gemfibrozil and its generated metabolites are shown in Figure 3. Crystallographic data and ORTEP diagrams of the asymmetric unit of the framework and analyte are shown in the Supplemental Data.



## Discussion

In summary, we were able to analyze the structure of three different phase I metabolites and one phase II acyl glucuronide of our tool compound gemfibrozil at atomic level with only ng amounts of substance using the crystalline sponge method. Data from MS analysis provided us information on the type of metabolism, but could not pinpoint the exact position of the hydroxylation, carboxylation or glucuronidation reaction. Only the crystallographic data showed that hydroxylation occurred at the 4'-C and 5'-CH<sub>3</sub> positions of the aromatic ring, as well as further oxidation at position 5'-CH<sub>3</sub> to form a carboxylic acid and direct glucuronidation at the carboxylic acid moiety of the alkyl chain. The Cambridge Crystallographic Data Centre did not contain crystallographic data of an acyl glucuronide before and could be extended through M4 gemfibrozil.

The need to identify the structure of metabolites meets the need of industry to assess the contribution of drug metabolites to efficacy, toxicity or drug-drug interactions. It also meets the requirements from regulatory agencies to ascertain that human metabolites are present in preclinical species used to assess drug safety. Especially acyl glucuronides play an important role in the assessment of drug-related risk factors, as they contain the risk to be potentially reactive metabolites. An intramolecular rearrangement process, so called "migration", leads to the formation of chemically reactive species which can covalently bind proteins (Zhong et al., 2015; Bailey and Dickinson, 2003).

The application of the crystalline sponge technology in structure elucidation is a convenient technology that can be applied in drug discovery to accelerate the metabolite identification processes and allow for using human-derived systems to identify human metabolites from in vitro incubation, years before clinical samples are available.

This technology is, however, in its infancy. Therefore optimization of soaking conditions (time, solvent and temperature) as well as assessing the purity required for compound is still needed. The amount of analyte per experiment was used in excess, but could be reduced as in the original publication mentioned (Inokuma et al., 2013). It has also been described that not all compounds can be applied in this technology as the size of the molecules may prevent them

from entering the pores and their polarity may prevent some molecules to dissolve in the hydrophobic solvents typically used during soaking. It was therefore unclear, if the technology would be applicable to drug metabolites resulting from phase I and especially phase II reactions. To answer this question, we used reference metabolites (M2 and M4 gemfibrozil) to confirm the ability of CS-XRD to assess the structure of phase I and phase II metabolites. Beside these 2 metabolites, we have applied this new technology to identify the structure of the 2 other unknown metabolites (M1 and M3 gemfibrozil) confirming its application in drug development of new chemical entities where only the structure of the candidate molecule is known.

Our results show that the combination of CS-XRD with MS data offers a great opportunity for scientists in drug metabolism and pharmacokinetics to assess the structure of metabolites produced in low amount from in vitro studies. The technology is also able to provide information on stereochemistry offering a significant advantage over other techniques like MS or NMR. In conclusion, we show that CS can be used to assess the structure of traces of drug-like molecules as well as their metabolites showing promising application at different stages of drug development where complete structure elucidation is needed. Besides metabolites, this technology could also be applied to identify degradation products and impurities during drug product development.

## **Acknowledgements**

The authors thank Andrea Unzue-Lopez for kindly providing 4'-OH gemfibrozil reference material, Axel-Walter Thomasberger, Ralf-Erwin Licht, Jonny Nachtigall and Birikiti Kidane for their kind technical support in the laboratory. The authors also thank Anna K. H. Hirsch for her constructive comments of the manuscript.

## **Authorship Contributions**

Participated in research design: all authors

Conducted experiments: Rosenberger, von Essen, Khutia

Contributed new reagents or analytic tools: Urbahns, Kühn

Performed data analysis: von Essen, Rosenberger

Wrote or contributed to the writing of the manuscript: Rosenberger, von Essen, Khutia, Kühn, Hartmann, Badolo

## **Conflict of interest**

The authors declare a conflict of interest. L.R., C.v.E., A.K., C.K., K.U., K.G. and L.B. are employees of Merck KGaA, Darmstadt, Germany.

## References

- Bailey MJ and Dickinson RG (2003) Acyl glucuronide reactivity in perspective: biological consequences. *Chem Biol Interact* 145(2): 117-137.
- Biradha K and Fujita M (2002) A springlike 3D-coordination network that shrinks or swells in a crystal-to-crystal manner upon guest removal or readsorption. *Angew Chem Int Ed Engl* 41(18): 3392–3395.
- Brandon EFA, Raap CD, Meijerman I, Beijnen JH, and Schellens JHM (2003) An update on in vitro test methods in human hepatic drug biotransformation research: pros and cons. *Toxicol Appl Pharmacol* 189(3): 233–246.
- Dolomanov OV, Bourhis LJ, Gildea RJ, Howard JAK, and Puschmann H (2009) OLEX2: a complete structure solution, refinement and analysis program. *J Appl Crystallogr* 42: 339-341.
- EMA (2013) Guideline on the investigation of drug interactions, European Medicines Agency, Amsterdam, Netherlands.
- Fasinu P, Bouic PJ, and Rosenkranz B (2012) Liver-based in vitro technologies for drug biotransformation studies - a review. *Curr Drug Metab* 13(2): 215-224.
- FDA (2016) Guidance for Industry: Safety testing of drug metabolites, US Department of Health and Human Services FDA, Center for Drug Evaluation and Research, Silver Spring, MD.
- Hoshino M, Khutia A, Xing H, Inokuma Y, and Fujita M (2016) The crystalline sponge method update. *IUCrJ* 3(2): 139-151.
- Inokuma Y, Ukegawa T, Hoshino M, and Fujita M (2016) Structure determination of microbial metabolites by the crystalline sponge method. *Chem Sci* 7:3910–3913.
- Inokuma Y, Yoshioka S, Ariyoshi J, Arai T, Hitora Y, Takada K, Matsunaga S, Rissanen K, and Fujita M (2013) X-ray analysis on the nanogram to microgram scale using porous complexes. *Nature* 495: 461-466; Corrigendum: Inokuma Y, Yoshioka S, Ariyoshi J, Arai T, Hitora Y, Takada K, Matsunaga S, Rissanen K, and Fujita M (2013) *Nature* 501: 262.

- Jones CG, Martynowycz MW, Hattne J, Fulton TJ, Stoltz BM, Rodriguez JA, Nelson HM, and Gonen T (2018) The CryoEM method MicroED as a powerful tool for small molecule structure determination. *ACS Cent Sci* 4(11): 1587–1592.
- Kai K, Sogame M, Sakurai F, Nasu N, and Fujita M (2018) Collimonins A–D, unstable polyynes with antifungal or pigmentation activities from the fungus-feeding bacterium collimonas fungivorans Ter331. *Org Lett* 20(12): 3536-3540.
- Kunde T and Schmidt BM (2019) Microcrystal electron diffraction (MicroED) for small- molecule structure determination. *Angew Chem Int Ed Engl* 58(3): 666-668.
- Ma S and Chowdhury SK (2011) Analytical strategies for assessment of human metabolites in preclinical safety testing. *Anal chem* 83(13): 5028-5036.
- Massa W (2004) *Crystal Structure Determination*, Springer-Verlag GmbH, Heidelberg.
- Morita I, Mori T, Mitsuhashi T, Hoshino S, Taniguchi Y, Kikuchi T, Nagae K, Nasu N, Fujita M, Ohwada T, and Abe I (2019) Exploiting a C-N bond forming cytochrome P450 monooxygenase for C-S bond formation. *Angew Chem Int Ed Engl* doi: 10.1002/anie.201916269.
- Murai T, Iwabuchi H, and Ikeda T (2004) Identification of gemfibrozil metabolites, produced as positional isomers in human liver microsomes, by on-line analyses using liquid chromatography/mass spectrometry and liquid chromatography/nuclear magnetic resonance spectroscopy. *J Mass Spectrom Soc Jpn* 52(5): 277-283.
- Nedderman AN (2009) Metabolites in safety testing: metabolite identification strategies in discovery and development. *Biopharm Drug Dispos* 30(4): 153–162.
- Prakash C, Shaffer CL, and Nederman A (2007) Analytical strategies for identifying drug metabolites. *Mass Spectrom Rev* 26(3): 340– 369.
- Prasad B, Garg A, Takwani H, and Singh S (2011) Metabolite identification by liquid chromatography-mass spectrometry. *Trends Analyt Chem* 30(2): 360-387.
- Ramadhar TR, Zheng SL, Chen YS, and Clardy J (2015) Analysis of rapidly synthesized guest-filled porous complexes with synchrotron radiation: practical guidelines for the crystalline sponge method. *Acta Crystallogr A Found Adv* 71(Pt 1): 46–58.

- Rigaku Oxford Diffraction (2018) *CrysAlis PRO*. Rigaku Corporation, Tokyo, Japan.
- Sakurai F, Khutia A, Kikuchi T, and Fujita M (2017) X-ray structure analysis of n-containing nucleophilic compounds by the crystalline sponge method. *Chem Eur J* 23(60): 15035-15040.
- Sheldrick GM (2015) Crystal structure refinement with SHELXL. *Acta Crystallogr C Struct Chem* 71: 3-8.
- Wada N, Kersten RD, Iwai T, Lee S, Sakurai S, Kikuchi T, Fujita D, Fujita M, and Weng JK (2018) Crystalline-sponge-based structural analysis of crude natural product extracts. *Angew Chem Int Ed Engl* 57(14): 3671-3675.
- Zhang D and Surapaneni S (2012) *ADME-Enabling Technologies in Drug Design and Development*, John Wiley & Sons, Inc., New Jersey.
- Zhang Z, Zhu M, and Tang W (2009) Metabolite identification and profiling in drug design: current practice and future directions. *Curr Pharm Des* 15(19): 2220-2235.
- Zhong S, Jones R, Lu W, Schadt S, and Ottaviani G (2015) A new rapid in vitro assay for assessing reactivity of acyl glucuronides. *Drug Metab Dispos* 43(11): 1711-1717.

## Figure Legends

**Figure 1.** Chemical structure of gemfibrozil

**Figure 2.** Workflow of the metabolite identification processes using crystalline sponge technology

**Figure 3.** Metabolic pathway of gemfibrozil; M1: 5'-CH<sub>2</sub>OH gemfibrozil; M2: 4'-OH gemfibrozil; M3: 5'-COOH gemfibrozil; M4: gemfibrozil glucuronide

**Figure 4.** Extracted ion chromatogram (XIC) of MRM for gemfibrozil with rat liver S9 (A) and human liver S9 (B) in the presence of NADPH and NADP<sup>+</sup> as cofactors. M1 (5'-CH<sub>2</sub>OH gemfibrozil): m/z 265.0/137.2, M2 (4'-OH gemfibrozil): m/z 265.0/137.2, M3 (5'-COOH gemfibrozil): m/z 279.0/151.0, gemfibrozil: m/z 249.1/121.0

**Figure 5.** (A) Asymmetric unit of [(ZnCl<sub>2</sub>)<sub>3</sub>(tpt)<sub>2</sub>]<sub>n</sub> with two 5'-hydroxymethyl gemfibrozil molecules and one cyclohexane. (B) Electron density map F<sub>o</sub> (contoured at 0.67σ (gemfibrozil), 0.65σ (M2 reference), 0.73σ (M1), 0.95σ (M2), 1.00σ (M3)) of benzene core from reference material and incubation samples

**Figure 6.** The crystal structure of M2 gemfibrozil exhibits π-π stacking interactions (grey line) and CH...Cl interactions (green dashed line) between the analyte and the framework. Centroids are shown as grey spheres in the center of the benzene ring.

Figure 1

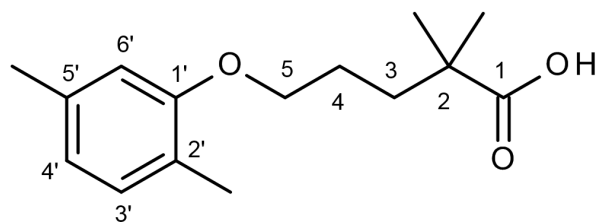




Figure 2

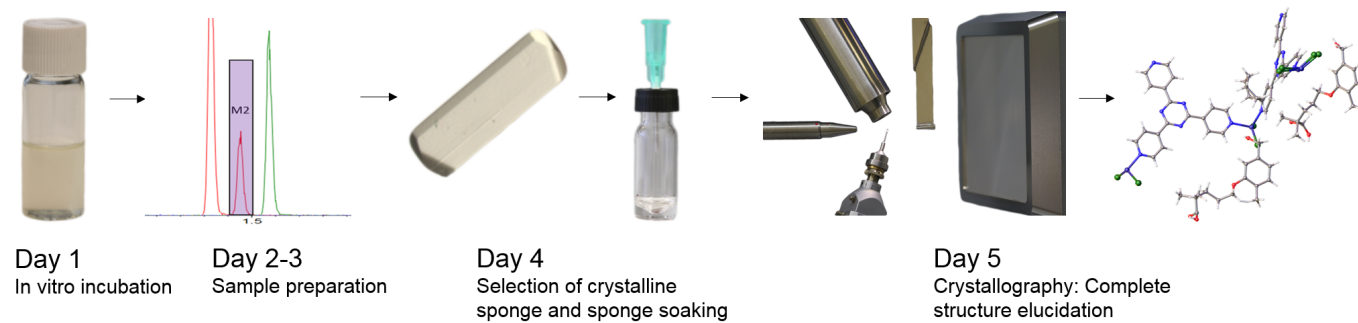


Figure 3

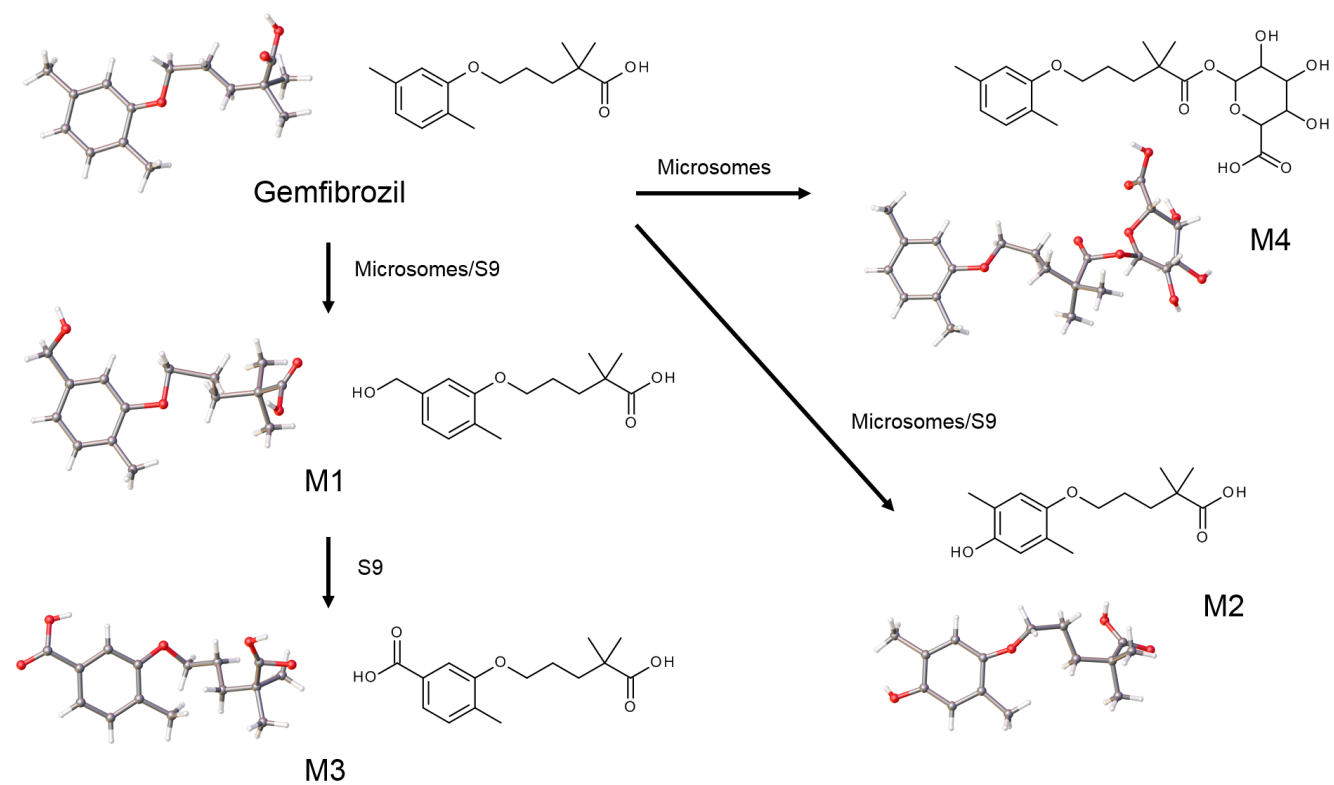


Figure 4

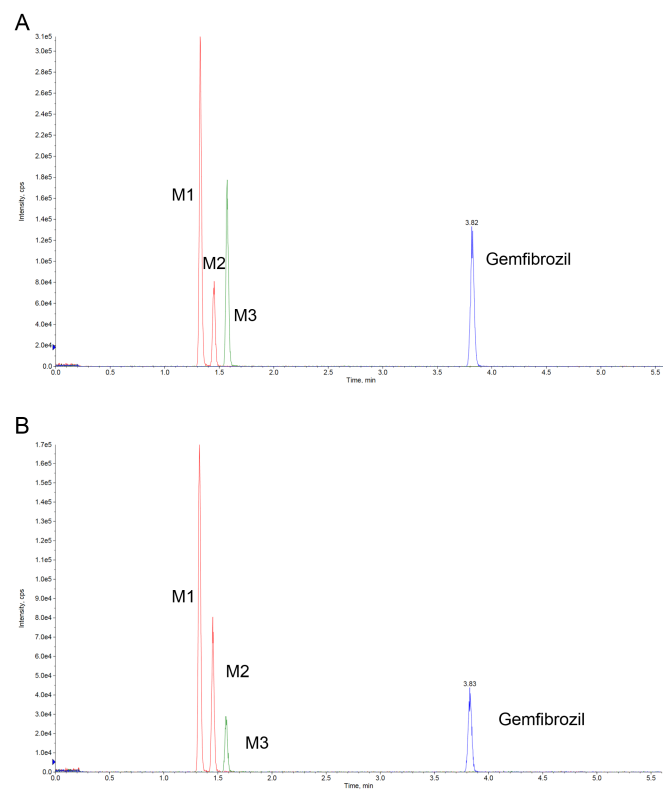
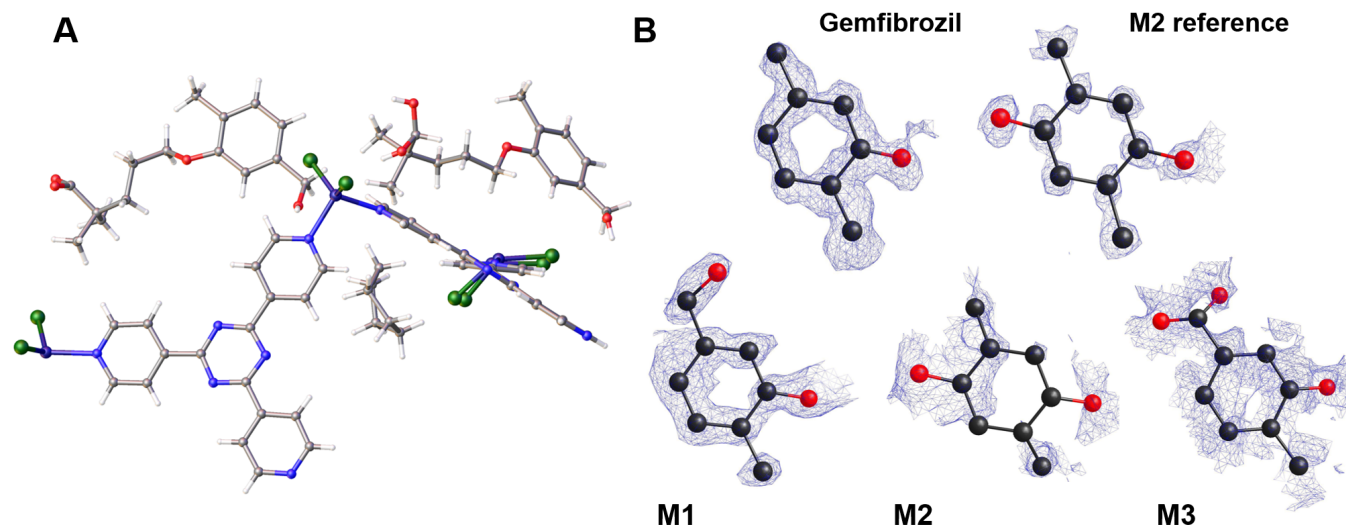


Figure 5



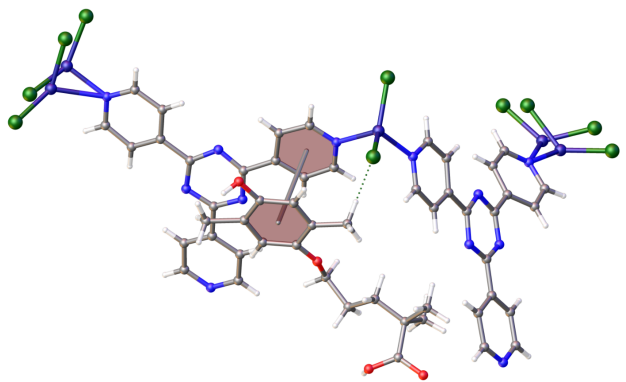


Figure 6

## **Supplemental Data for**

### **Crystalline Sponges as a sensitive and fast method for metabolite identification: Application to gemfibrozil and its phase I and II metabolites**

Lara Rosenberger, Carolina von Essen, Anupam Khutia, Clemens Kühn, Klaus Urbahns,  
Katrini Georgi, Rolf W. Hartmann, Lassina Badolo

Department of Drug Metabolism and Pharmacokinetics (DMPK), Merck KGaA, Frankfurter  
Strasse 250, 64293 Darmstadt, Germany (L.R., K.U., K.G., L.B.)

Innovation Center, Merck KGaA, Frankfurter Strasse 250, 64293 Darmstadt, Germany (C.v.E.,  
A.K., C.K.)

Department of Drug Design and Optimization (DDOP), Helmholtz-Institute for Pharmaceutical  
Research Saarland (HIPS)-Helmholtz Centre for Infection Research (HZI), Campus E8.1,  
66123 Saarbrücken, Germany (L.R., R.W.H.)

## **Supplemental Methods**

### **Single crystal X-ray diffraction experiments**

Single crystal X-ray diffraction measurements were conducted on a Rigaku Oxford Diffraction XtaLAB Synergy-R diffractometer using Cu-K $\alpha$  X-ray radiation ( $\lambda = 1.54184 \text{ \AA}$ ), equipped with a HyPix-6000HE (M1 gemfibrozil, M2 gemfibrozil, M3 gemfibrozil and M4 gemfibrozil) or HyPix-Arc 150° (gemfibrozil) Hybrid Photon Counting (HPC) detector (Rigaku, Tokyo, Japan) at a temperature of 100 K using a Cryostream 800 nitrogen stream (Oxford Cryostreams, UK). The software CrysAlisPro ver. 171.39.46 (M1 gemfibrozil, M2 gemfibrozil, M3 gemfibrozil and M4 gemfibrozil) and ver. 171.41.48a (gemfibrozil) was used for calculation of measurement strategy and data reduction (data integration, empirical and numerical absorption corrections and scaling). The software ShelXle was used to generate the electron density maps  $F_o$  to verify the position of metabolism (Hübschle et al., 2011). Isocontour levels are described as  $\sigma$  (square root of the average variance of the density).

### **Crystal structure analysis**

All crystal structures were modeled using OLEX2 (Dolomanov et al., 2009), solved with SHELXT ver. 2014/5 and refined using SHELXL ver. 2018/1 (Sheldrick, 2015). Figures of framework and analyte were created using OLEX2. Non-hydrogen atoms were refined anisotropically. Hydrogen atoms were fixed using the riding model. Populations of the guests in the crystal were modelled by least-square refinement of a guest/solvent disorder model under the constraint that the sum of them should equal to 100%. When a reasonable guest/solvent disorder model could not be built due to severe disorder of the solvent molecules, the guest populations were set to 100% (s.o.f. = 1). In this case, overestimation is unavoidable if the actual occupancy is less than 100% leading to high large displacement parameters compared to the framework (Hoshino et al., 2016). Number of restraints was tried to minimize and applied without changing the standard deviation. Solvent cyclohexane molecules in the pores were found in the difference electron density map and refined using the restraints (DFIX,

DANG, SIMU and ISOR). These molecules are expected to be severely disordered due their high thermal motion. Due to their averaged structure of various geometry and orientation, some cyclohexane molecules are distorted to energetically-unfavorable (boat-shaped or twisted) structures. Some "Alert A" notifications were found in the validation program CheckCIF. Those alerts are derived from short intermolecular contacts of hydrogen atoms between solvent molecule (cyclohexane) and the crystalline sponge framework, guest molecule, and other solvent molecules, and are unavoidable due to severe disorder. The comments for the alerts are described in the CIFs using the validation response form (vrf).

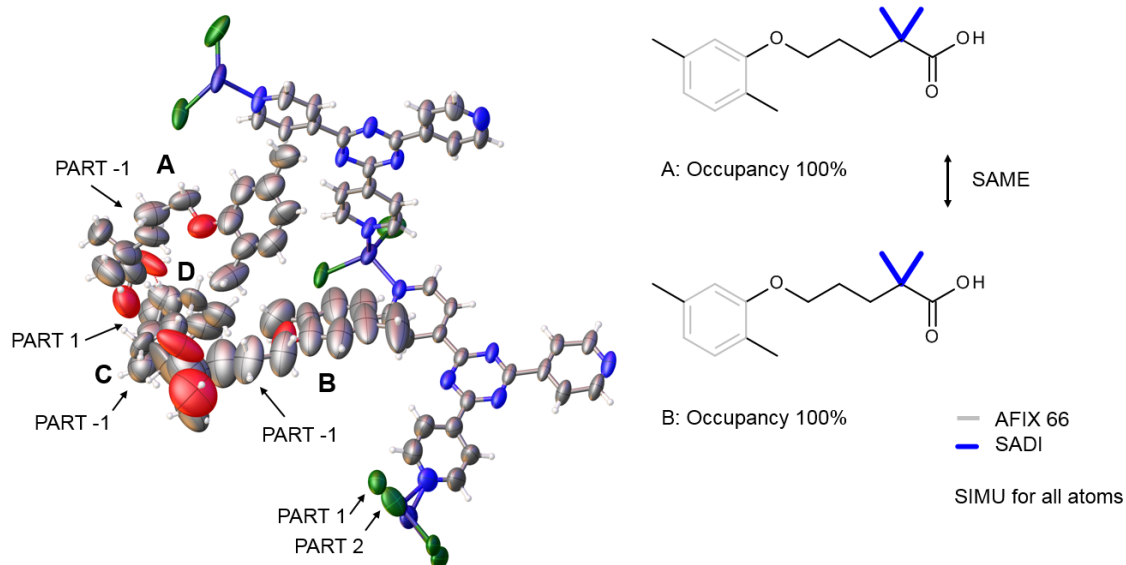
The present crystal structures are not used for exact structure analysis. They are used to confirm structure proposals derived from information obtained from mass spectrometry and knowledge of the parent structure. Therefore, details of the crystals structure (bond lengths, angles, etc.) will not be discussed.



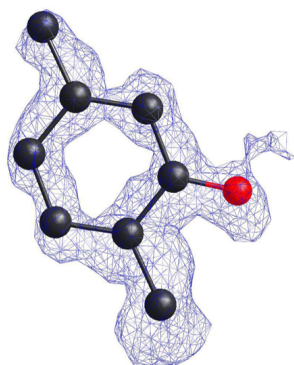
## Crystallographic data for gemfibrozil

Crystal size:  $131 \times 83 \times 59 \mu\text{m}^3$ , refined formula:  $\text{C}_{55}\text{H}_{55}\text{Cl}_6\text{N}_{12}\text{O}_3\text{Zn}_3$ , formula weight ( $M_r$ ) = 1344.24 g/mol, yellow needle, crystal system: Monoclinic, space group  $C2/c$ ,  $Z = 8$ , 14170 unique reflections merged from recorded 47910 ones ( $2.695^\circ < \theta < 71.380^\circ$ ) were used for structural analysis ( $R_{\text{int}} = 0.0181$ ). Lattice parameters, R-factor on  $F^2 > 2\sigma(F^2)$ , weighted R-factor, and goodness-of-fit are follows:  $a = 33.6364(5) \text{ \AA}$ ,  $b = 14.36600(10) \text{ \AA}$ ,  $c = 32.0319(4) \text{ \AA}$ ,  $\beta = 102.8500(10)^\circ$ ,  $V = 15090.8(3) \text{ \AA}^3$ ,  $R = 0.1125$ ,  $wR_2 = 0.3255$ ,  $S = 1.102$ . Calculated density is  $1.183 \text{ g cm}^{-3}$ . Linear absorption coefficient ( $\mu$ ) is  $3.409 \text{ mm}^{-1}$ . Residual electron density (max/min) is  $0.865/-0.471 \text{ e \AA}^{-3}$ . CCDC number 1983846. The ORTEP diagram of the asymmetric unit of the framework and gemfibrozil is shown in Supplemental Figure 1.

The framework is refined without using restraints. One  $\text{ZnCl}_2$  moiety is disordered and refined using disorder model. Two guest molecules (A, B) were found in the asymmetric unit. Two additional cyclohexane molecules (C, D) can be found in the asymmetric unit. Both analytes are located very close to the symmetry centers (inversion and  $C2$  axis, respectively) and are refined using PART -1 command to avoid calculation of bonds between symmetry related atoms. The analyte occupancy is 100%. The geometry of both molecules is related by SAME command. The benzene ring was fixed using AFIX 66. Two C-C single bonds were restrained using SADI command. No other geometrical restrains were used in the refinement of the analyte molecules. SIMU was applied for all analyte molecules. Applied restraints can be taken from Supplemental Fig. 1. Supplemental Fig. 2 shows the electron density map of the aromatic part of one gemfibrozil molecule taken from the crystal structure.



Supplemental Fig. 1. ORTEP diagram with 50% probability of the asymmetric unit of gemfibrozil; restraints and constraints applied in the refinement of gemfibrozil



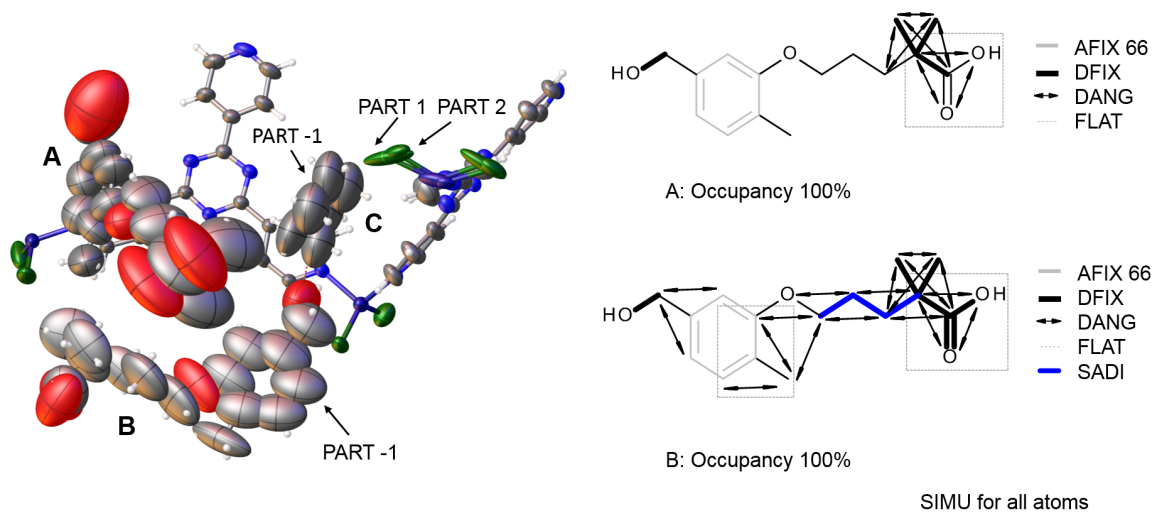
Supplemental Fig. 2. Electron density map  $F_o$  of the benzene core from gemfibrozil (contoured at the  $0.67\sigma$  level)

## Crystallographic data for M1 gemfibrozil

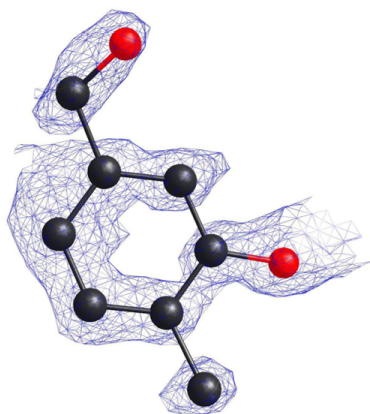
Crystal size:  $163 \times 76 \times 54 \mu\text{m}^3$ , refined formula:  $\text{C}_{62}\text{H}_{63}\text{Cl}_6\text{N}_{12}\text{O}_6\text{Zn}_3$ , formula weight ( $M_r$ ) = 1475.05  $\text{g/mol}$ , yellow needle, crystal system: Monoclinic, space group  $\text{C2/c}$ ,  $Z = 8$ , 13144 unique reflections merged from recorded 42219 ones ( $2.692^\circ < \theta < 67.080^\circ$ ) were used for structural analysis ( $R_{\text{int}} = 0.0218$ ). Lattice parameters, R-factor on  $F^2 > 2\sigma(F^2)$ , weighted R-factor and goodness-of-fit are follows:  $a = 33.5373(8) \text{ \AA}$ ,  $b = 14.3660(2) \text{ \AA}$ ,  $c = 31.4546(9) \text{ \AA}$ ,  $\beta = 101.801(3)^\circ$ ,  $V = 14834.4(6) \text{ \AA}^3$ ,  $R = 0.0758$ ,  $wR_2 = 0.2571$ ,  $S = 1.056$ . Calculated density is  $1.321 \text{ gcm}^{-3}$ . Linear absorption coefficient ( $\mu$ ) is 3.547. Residual electron density (max/min) is 0.703/-0.622  $\text{e\AA}^{-3}$ . CCDC number 1983847. The ORTEP diagram of the asymmetric unit of the framework and M1 gemfibrozil is shown in Supplemental Fig. 3.

The framework is refined without using restraints. One  $\text{ZnCl}_2$  moiety is disordered and refined using disorder model. Two guest molecules were found in the asymmetric unit. One molecule (A) is ordered at general position with modeled occupancy of 100% because solvent disorder models could not be built due to high disorder. The other molecule (B) is disordered at special position ( $\text{C2}$  axis) and refined using PART -1 command and s.o.f. = 0.5. One cyclohexane molecule (C) is disordered on special position ( $\text{C2}$  axis) and refined using PART -1 command and s.o.f. = 0.5. Some restraints were applied for refinement of the analyte and solvent molecules. Both guest molecules were refined with applying SIMU (for complete molecules). The restraints used for the refinement are summarized in Supplemental Fig. 3. The benzene part of both molecules could be refined with minor restraints and the presence of oxidation at the 5'-methyl carbon is clearly visible from electron density maps (Supplemental Fig. 4). The positions of carbon atoms in the alkyl chains had to be restrained to their theoretical values applying DFIX and DANG commands due to high mobility and not resolved disorder with solvent molecules of these parts. In the present study, the task of the crystalline sponge method was the determination of positions of oxidation during metabolization and the position could be determined without any doubts. The combination of knowledge of the parent structure, information from mass spectrometry and crystalline sponge analysis leads to a successful structure elucidation. For a full structural analysis of this compound without any prior

knowledge of the composition, soaking conditions should be optimized for higher occupancy leading to improved data.



Supplemental Fig. 3. ORTEP diagram with 50% probability of the asymmetric unit of M1 gemfibrozil; restraints and constraints applied in the refinement of M1 gemfibrozil

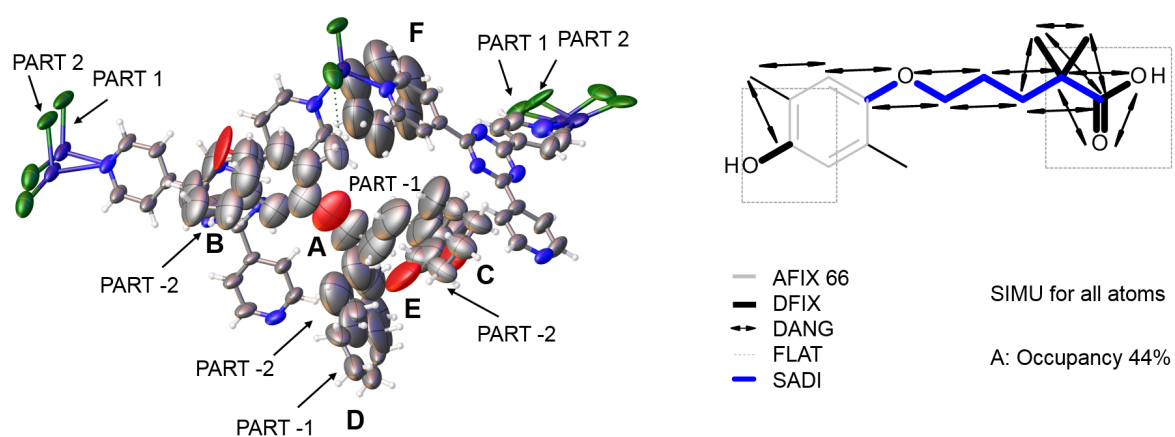


Supplemental Fig. 4. Electron density map  $F_o$  of the benzene core from M1 gemfibrozil (contoured at the 0.73 $\sigma$  level)

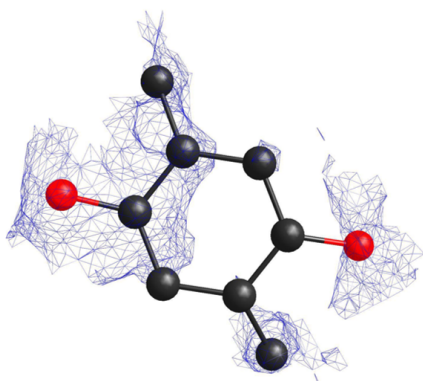
## Crystallographic data for M2 gemfibrozil

Crystal size:  $295 \times 144 \times 120 \mu\text{m}^3$ , refined formula:  $\text{C}_{52}\text{H}_{53}\text{Cl}_6\text{N}_{12}\text{O}_1\text{Zn}_3$ , formula weight ( $M_r$ ) = 1262.43  $\text{g/mol}$ , light yellow needle, crystal system: Monoclinic, space group  $C2/c$ ,  $Z = 8$ , 12268 unique reflections merged from recorded 38793 ones ( $3.373^\circ < \theta < 67.076^\circ$ ) were used for structural analysis ( $R_{\text{int}} = 0.0209$ ). Lattice parameters, R-factor on  $F^2 > 2\sigma(F^2)$ , weighted R-factor, and goodness-of-fit are follows:  $a = 31.8982(5) \text{ \AA}$ ,  $b = 14.4063(2) \text{ \AA}$ ,  $c = 30.4802(6) \text{ \AA}$ ,  $\beta = 98.783(2)^\circ$ ,  $V = 13842.5(4) \text{ \AA}^3$ ,  $R = 0.0756$ ,  $wR_2 = 0.2381$ ,  $S = 1.084$ . Calculated density is  $1.212 \text{ g cm}^{-3}$ . Linear absorption coefficient ( $\mu$ ) is  $3.658 \text{ mm}^{-1}$ . Residual electron density (max/min) is  $0.644/-0.610 \text{ e \AA}^{-3}$ . CCDC number 1983848. The ORTEP diagram of the asymmetric unit of the framework and M2 gemfibrozil is shown in Supplemental Fig. 5.

The framework exhibits disorder of two  $\text{ZnCl}_2$  parts. In one case, displacement parameters of respective atoms were restrained using the DELU command. One guest molecule (A) can be found in the asymmetric unit disordered with two cyclohexane (B, C) molecules. Two additional cyclohexane molecules, one disordered with a second cyclohexane (D, E) and one without disorder (F) can be found in the asymmetric unit. The analyte and cyclohexane molecules lie near the center of symmetry ( $C2$  axis). Therefore, the analyte is refined using PART -1 with s.o.f. =  $0.5 \times (\text{FVAR}\#4)$  and the disordered cyclohexane molecules are refined using PART -2 with s.o.f. =  $0.5 \times (1 - \text{FVAR}\#4)$ . This leads to an analyte occupancy of 44%. AFIX 66 was applied to the benzene ring, and DFIX and DANG were used for some bonds and angles of the guest (see Supplemental Fig. 5). Due to missing hydrogen bond acceptors near phenyl and carboxylic acid alcohol function and low occupancy, a stable refinement for the respective hydrogen atom was not possible. SIMU was applied for both analyte molecules. The oxidation of the benzene ring in 4-position during metabolization is clearly visible from electron density maps (Supplemental Fig. 6).



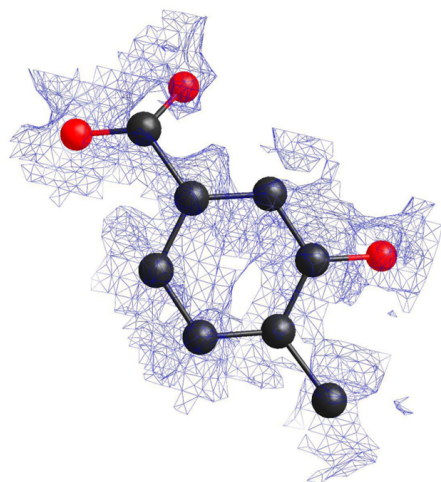
Supplemental Fig. 5. ORTEP diagram with 50% probability of the asymmetric unit of M2 gemfibrozil; restraints and constraints applied in the refinement of M2 gemfibrozil



Supplemental Fig. 6. Electron density map  $F_o$  of the benzene core from M2 gemfibrozil (contoured at the  $0.95\sigma$  level)

### Crystallographic data for M3 gemfibrozil

The framework can be refined without using restraints. One disordered  $\text{ZnCl}_2$  moiety is modeled using the disorder model. Applying solvent masking implemented in OLEX2 after complete framework refinement, one void with  $V = 969 \text{ \AA}^3$  and 264 electrons per asymmetric unit can be found. This is consistent with the presence of 1.5 molecules of M3 and one cyclohexane molecule (273 electrons). Electron density assignable to a benzene core can be found. Careful investigation of the electron density map around the benzene core revealed electron density assignable to a carboxylic acid function in 5 position (Supplemental Fig. 7). Due to low occupancy and high disorder, the complete analyte molecule could not be modeled. Information obtained from MS indicated a two-fold oxidation, presumably to a carboxylic acid and the structure obtained from the crystalline sponge method confirmed the position of oxidation.



Supplemental Fig. 7. Electron density map  $F_o$  of the benzene core from M3 gemfibrozil (contoured at the  $1.00\sigma$  level)

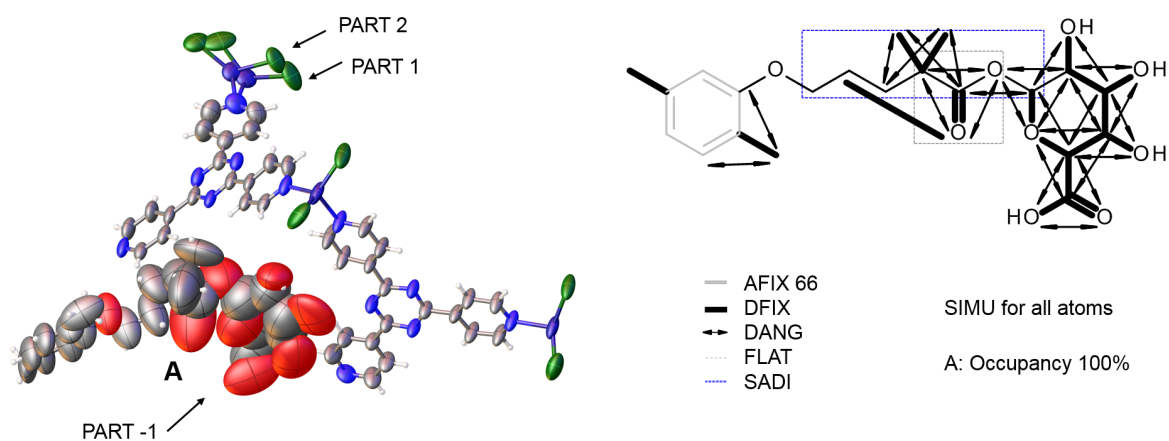
## Crystallographic data for M4 gemfibrozil

Crystal size:  $200 \times 97 \times 26 \mu\text{m}^3$ , refined formula:  $\text{C}_{47}\text{H}_{39}\text{Cl}_6\text{N}_{12}\text{O}_5\text{Zn}_3$ , formula weight ( $M_r$ ) = 1246.70  $\text{g/mol}$ , colourless needle, crystal system: Monoclinic, space group  $C2/c$ ,  $Z = 8$ , 12853 unique reflections merged from recorded 39917 ones ( $2.742^\circ < \theta < 67.079^\circ$ ) were used for structural analysis ( $R_{\text{int}} = 0.0187$ ). Lattice parameters, R-factor on  $F^2 > 2\sigma(F^2)$ , weighted R-factor, and goodness-of-fit are follows:  $a = 32.8136(6) \text{ \AA}$ ,  $b = 14.42330(10) \text{ \AA}$ ,  $c = 31.0646(5) \text{ \AA}$ ,  $\beta = 100.840(2)^\circ$ ,  $V = 14439.9(4) \text{ \AA}^3$ ,  $R = 0.0699$ ,  $wR_2 = 0.2481$ ,  $S = 1.026$ . Calculated density is  $1.147 \text{ gcm}^{-3}$ . Linear absorption coefficient ( $\mu$ ) is  $3.544 \text{ mm}^{-1}$ . Residual electron density (max/min) is  $0.724/-0.460 \text{ e\AA}^{-3}$ . CCDC number 1983849. The ORTEP diagram of the asymmetric unit of the framework and M4 gemfibrozil is shown in Supplemental Fig. 8.

The framework exhibits disorder of one  $\text{ZnCl}_2$  part refined using the disorder model. One guest molecule (A) can be found in the asymmetric unit near the center of symmetry ( $C2$  axis) and refined using PART -1 command with s.o.f. = 0.5. The analyte occupancy was modeled to 100% as no disordered solvent molecules could be assigned. AFIX 66 was applied to the benzene ring, and DFIX and DANG were used for some bonds and angles of the guest (see Supplemental Fig. 8). The guest molecule was refined with applying SIMU. Additional electron density was found that could not be assigned due to disorder. Contribution of this electron density was removed using solvent masking algorithm implemented in OLEX2. A solvent mask was calculated and 648 electrons were found in a volume of  $2548 \text{ \AA}^3$  in one void per unit cell. This is consistent with the presence of 1.5 cyclohexane molecules per asymmetric unit which account for 576 electrons per unit cell. The analyte was used as enantiopure material. Although it is in general possible to determine absolute structures with the crystalline sponge method, we were not able to determine the absolute structure of gemfibrozil glucuronide. Due to disorder, conformation of the glucuronide, i. e. oxane and substituents cannot be determined. The presence of aliphatic rings in unfavorable conformations is often observed in the crystalline sponge method especially for cyclohexane molecules. Nevertheless, the position of glucuronidation at the carboxylic acid function of gemfibrozil can be confirmed.



Material obtained from incubation was used for a soaking experiment. The crystal structure shows electron density assignable to gemfibrozil glucuronide. Due to high disorder and low analyte occupancy, a stable model could not be obtained. Therefore, the structure obtained from reference material is presented.



Supplemental Fig. 8. ORTEP diagram with 50% probability of the asymmetric unit of M4 gemfibrozil; restraints and constraints applied in the refinement of M4 gemfibrozil

## References

- Dolomanov OV, Bourhis LJ, Gildea RJ, Howard JAK, and Puschmann H (2009) *OLEX2*: a complete structure solution, refinement and analysis program. *J Appl Crystallogr* 42: 339-341.
- Hübschle CB, Sheldrick GM, and Dittrich B (2011) *ShelXle*: a Qt graphical user interface for SHELXL. *J Appl Crystallogr* 44(Pt 6): 1281-1284.
- Sheldrick GM (2015) Crystal structure refinement with SHELXL. *Acta Crystallogr C Struct Chem* 71: 3-8.