ORGANIC PROCESS RESEARCH & DEVELOPMENT
OPR&D

Open Access
This article is licensed under CC-BY 4.0

pubs.acs.org/OPRD
Article

Benchtop NMR-Based In-Line Analysis of Diastereoselective Enzymatic α-Amino Acid Synthesis: Quantification and Validation

Luca F. Schmidt, Logia Jolly, Leon Hennecke, Fernando Lopez Haro, Harald Gröger, and Andreas Liese*

Cite This: Org. Process Res. Dev. 2024, 28, 3791−3800
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: This study investigates the application of a commercial low-field benchtop NMR for real-time monitoring of enzymatically catalyzed reactions, focusing on the diastereoselectivity of the threonine aldolase-catalyzed stereoselective aldol reaction between glycine and benzaldehyde. Despite the signal overlap inherent in the weak electromagnetic field of the benchtop NMR system, a complemental hard modeling (CHM) approach effectively differentiates between diastereomers, enabling the determination of enzymatic diastereoselectivity and the transition from kinetic to thermodynamic control. In particular, the achievement of thermodynamic equilibrium in the enzymatic aldol reaction is observed for the first time using in-line methods, occurring at 30% benzaldehyde conversion after 2 h. In-line NMR analysis reveals a diastereomeric excess of 37:63 (*erythro*/*threo*), which closely aligns with off-line measurements via GC and HPLC (36:64). This determination of diastereomers using CHM enhances the efficiency of in-line monitoring in enzymatic reactions, promising significant advancements in pharmaceutical process development. Overall, the study underscores the utility of benchtop NMR systems for in-line analysis of enzymatic reactions, offering insights into reaction mechanisms, selectivity, and equilibrium dynamics, thereby facilitating more efficient process optimization in the area of fine chemicals.

KEYWORDS: *NMR spectroscopy, biocatalysis, analytics, complemental hard modeling, enzymatic reaction*

Downloaded via TU HAMBURG-HARBURG on November 28, 2024 at 12:49:56 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

1. INTRODUCTION

Due to the molecular structure-dependent mode of action of many pharmaceutical substances, diastereomeric purity is of absolute necessity. Since many pharmaceutical substances have a mode of action that depends on their molecular structure, stereoisomeric, e.g., diastereomeric, purity is crucial.[1] However, caused by the physical and chemical similarities of chiral molecules, the process of separating them into individual isomers can be resource- and cost-intensive. To circumvent separation processes, methods for the selective preparation of chiral molecules have been developed. Nevertheless, these chemical processes inevitably involve additional costs due to selective synthesis with several protection and deprotection steps, depending on the complexity of the target molecule.[2]

One major advantage of biocatalytic processes is their ability to facilitate selective reaction catalysis in fewer steps. This is often achieved with less harmful chemicals and requiring less energy input in comparison to classical chemical systems, which enables the design of more sustainable and simpler syntheses.[3]

Threonine aldolases are a particularly interesting and promising tool for synthesizing β-hydroxy-α-amino acids in an enantio- and diastereoselective fashion. Threonine aldolases are able to catalyze the C−C coupling of amino acids (e.g., serine or glycine) with an aldehyde to form a β-hydroxy α-amino acid with two stereogenic centers.[4−10] This approach enables the synthesis of a diverse range of important products. Diastereomeric pure aromatic sphingosines can be produced as the basic building block for antibiotics (thiamphenicol and vancomycin) or anti-inflammatories (cyclomarins).[11] Challenges exist in syntheses catalyzed by threonine aldolases, which give a poor selectivity at the β-carbon.[12]

The coupling of glycine and benzaldehyde results in the formation of two diastereomers with high selectivity at the α-carbon and, depending on the enzyme specification, low to medium selectivity for the β-carbon. The reaction system is seen in Figure 1. Aldol reactions occur in a kinetically

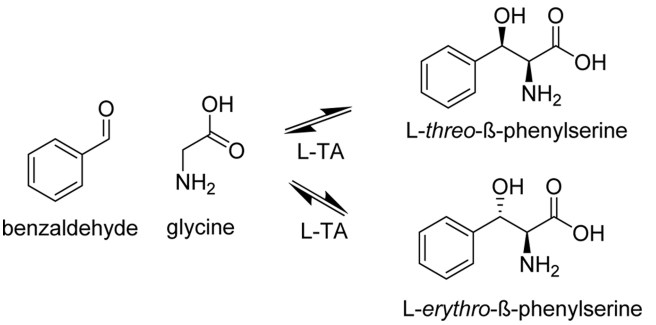
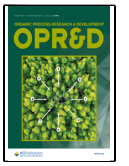


Figure 1. Reversible reaction of benzaldehyde and glycine to L-*threo*-β-phenylserine and L-*erythro*-β-phenylserine, catalyzed by L-threonine aldolase.

Received: February 28, 2024
Revised: July 19, 2024
Accepted: August 5, 2024
Published: September 13, 2024



ACS Publications
© 2024 The Authors. Published by American Chemical Society

https://doi.org/10.1021/acs.oprd.4c00076
Org. Process Res. Dev. 2024, 28, 3791−3800

controlled manner within the initial phase of the reaction, with thermodynamic equilibrium being reached later in the reaction course. This reaction behavior has been reported in the literature. Kimura et al. showed a diastereomeric excess (de) value of 40:60 (*erythro*/*threo*), while Gutierrez et al. showed 37:63 (*erythro*/*threo*).[12−14]

The similarities in the properties of diastereomers make them analytically challenging to distinguish, which can pose difficulties in reaction monitoring. The use of off-line analytical methods like gas chromatography (GC), high-performance liquid chromatography (HPLC), and high-field nuclear magnetic resonance spectroscopy (NMR) is typical for diastereomer differentiation; however, off-line measurement methods often cause a time lag, making it challenging to accurately determine diastereomer concentrations.[15]

NMR spectroscopy is capable of identifying the chemical environment of atoms with an odd number of protons and neutrons in the nucleus, which facilitates not only the identification of molecules but also their structural analysis.[16] It is imperative to acknowledge that high-field NMR systems do find application in industrial settings. However, their restricted usage is primarily as analytical tools for in-lab reaction monitoring rather than for extensive industrial-scale analysis.[17]

Benchtop NMR systems are increasingly being used in the pharmaceutical industry for the analysis of diastereomeric α-amino acids, due to their ability to provide high-resolution NMR spectra and their cost-effectiveness compared to high-field NMR systems.[18]

The main difference between high-field NMR and benchtop NMR spectrometers is their magnetic field strength. High-field NMR spectrometers typically have a magnetic field strength of 300 MHz or higher, whereas benchtop NMR spectrometers usually have magnetic field strengths of up to 100 MHz. The strength of the magnetic field can significantly affect NMR measurements. In high-field NMR systems, the discrepancy between the Larmor precession is notably larger compared to the coupling constants, where $\Delta\nu \gg J$. These spectra are in the so-called first order. As the Larmor precession approaches the coupling constant, the spectra become second order. With a decrease in the magnetic field strength, the Larmor precession increases, while the coupling constant remains unaffected. Consequently, a spectrum that would typically be considered first order in a high-field NMR system might appear as a second order spectrum in a low-field system. Differences observed at various magnetic field strengths can include varying peak intensities and numbers of peaks.[19] In terms of cost, high-field NMR spectrometers are significantly more expensive than low-field NMR systems. The reason for this is that high-field NMR spectrometers require a potent superconducting magnet and specialized infrastructure to operate. In contrast, low-field NMR systems are more affordable and accessible, making them a practical choice for routine analysis, process monitoring, and quality control by flow through analysis.[20−24]

This article focuses on the discussion of in-line monitoring, utilizing benchtop NMR, of the progress of the enzymatically catalyzed aldol reaction of glycine and benzaldehyde under the formation of L-*erythro*-β-phenylserine and L-*threo*-β-phenylserine by the enzyme threonine aldolases.

1.1. Indirect Hard Modeling. Statistical methods, such as partial least squares struggle to determine concentrations in reactive mixtures with complex molecular interactions and overlapping bands. In response to these challenges, indirect hard modeling (IHM) has emerged as a promising solution. IHM bridges the gap between soft and hard modeling techniques, providing a more comprehensive approach to spectral analysis.[25]

In IHM, pure component spectra are modeled as combinations of individual bands, allowing the inclusion of complex spectral features inherent in reactive mixtures. Mixture spectra are then expressed as weighted combinations of these pure component models, which again are sums of parametrized peak functions (Gaussian−Lorentzian). What sets IHM apart is its ability to account for nonlinear effects by allowing limited variation of spectral parameters. This feature makes IHM particularly useful in scenarios where conventional methods struggle due to the presence of reactive intermediates or unknown species.[26]

While IHM requires manual effort to model individual bands, its ability to handle strong intermolecular interactions and overlapping bands makes it a valuable tool in analytical chemistry. In addition, its adaptability to nonlinear effects broadens its applicability to diverse chemical systems, offering new ways to accurately determine concentrations in complex environments. In addition, it is possible to determine the intersecting areas in the event of strongly overlapping peaks. Since L-*erythro*-β-phenylserine is a rather volatile molecule and not available commercially, a complemental hard modeling (CHM) approach is used. CHM is well-suited for identifying an unknown spectrum within a fully known mixture. In this approach, a single mixture spectrum is sufficient as an input for the model.[27]

2. EXPERIMENTAL SECTION

2.1. Enzyme Synthesis. In this study, the L-threonine aldolases derived from the *Escherichia coli* strains ME9012 and GS245 were produced in BL21 *E. coli* cultures. The expression vector pET28a wt-ltaE(N)6His was employed to facilitate the production of the aldolases. A 20 mL preculture was prepared and subsequently incubated for 4 h at 37 °C in LB medium (Carl Roth, Karlsruhe, Germany). The selection of suitable cultures was achieved by employing kanamycin resistance mediated by the plasmid.[28] Detailed information regarding the protein and gene sequences can be found in the Supporting Information section.

The optical density of the preculture was determined to be 600 nm. Once the optical density reached 3, the preculture was transferred to the main culture with a total volume of 200 mL. After an additional 2 h of incubation, when the optical density was approximately 1.5 (0.6 to 0.8), enzyme synthesis was induced using isopropyl β-D-1-thiogalactopyranoside (IPTG). The induction was carried out for 16 h at 16 °C.

Once the induction was completed, the cells were purified, and ultrasonic cell disruption was carried out. The Sonopuls Generator GM 2070 (Bandelin GmbH, Berlin, Germany) was employed for this purpose, with the following settings: 3 cycles of 3 min each at 9 × 10 and 70% power. The ultrasound probe used had a diameter of 3 mm. The resulting solution was transferred to 1.5 mL reaction tubes and centrifuged for 30 min at 15,000 rpm and 4 °C. The supernatant was collected as the desired cell-free extract (CFE).

2.2. Experiment Setup. *2.2.1. In-Line Measurements.* The reaction medium consisted of 1 mol/L glycine (Merck Germany) (1), 100 mmol/L benzaldehyde (Merck Germany) (2), and 50 μmol/L pyridoxal-5′-phosphate (PLP) in a

https://doi.org/10.1021/acs.oprd.4c00076
Org. Process Res. Dev. 2024, 28, 3791−3800

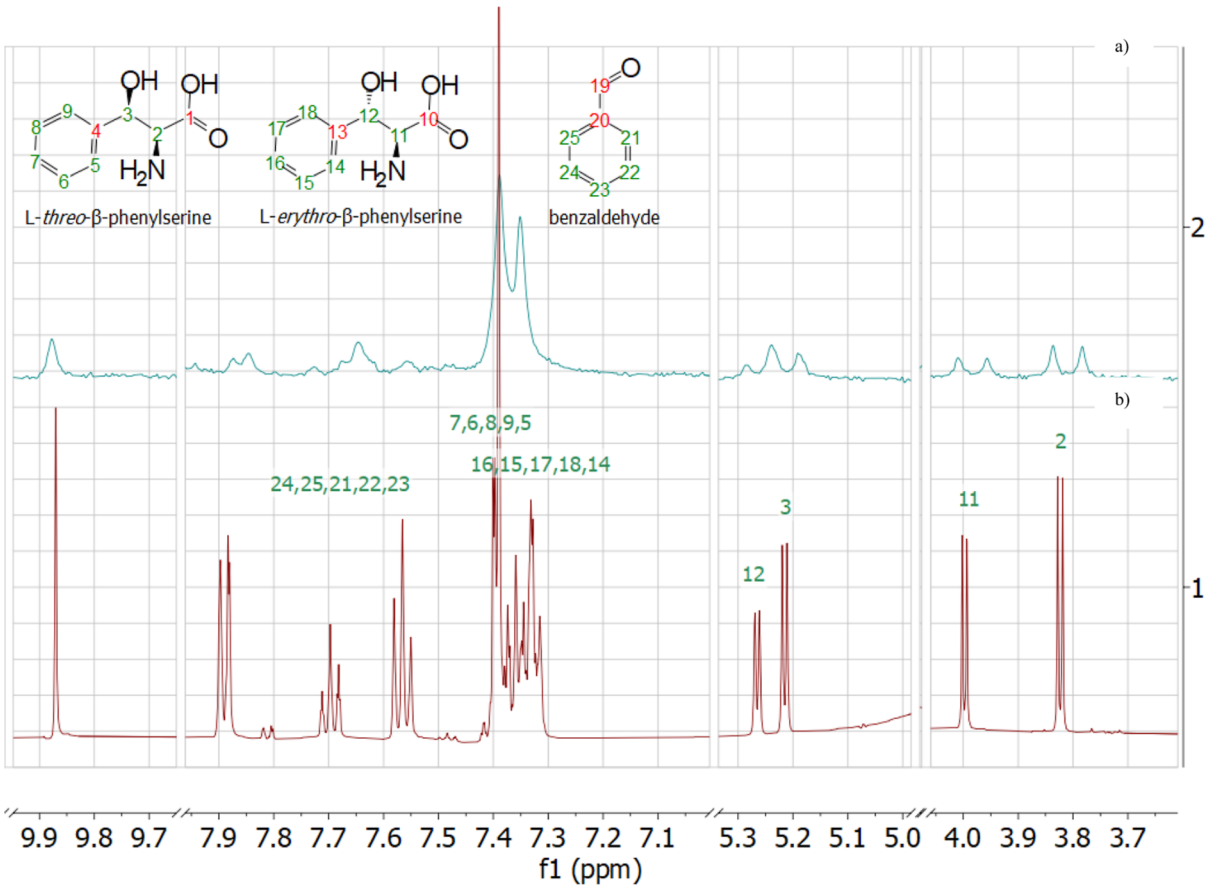


Figure 2. ^{1}H NMR Spectrum of the aldol reaction. Starting conditions: 1 M glycine; 100 mM benzaldehyde in 20 v/v DMSO in KPi buffer pH 8.0 V = 30 mL; T = 30 °C; sample taken after 5 h reaction time and (a) measured by low-field NMR Spinsolve 80 MHz ULTRA (Magritek, Aachen) (b) measured by Bruker Avance I 600 MHz (AV700).

phosphate buffer (Merck Germany) pH 8 with 20 v/v % dimethyl sulfoxide (DMSO) (Merck Germany) in a total volume of 30 mL. To avoid oxidation of the substrate benzaldehyde to benzoic acid, the entire reaction is conducted under an inert argon atmosphere. NMR measurements were performed with a Spinsolve 80 Carbon Ultra (Magritek GmbH). The software used for reaction monitoring is Spinsolve 2.0.2 (Magritek) in RMX mode. A shimming procedure was performed after each measured spectrum to ensure a homogeneous electromagnetic field and to reduce the noise in the measured spectra.

Following the first shimming and measurement of the spectrum, 1 mL of the CFE was injected into the reaction system, which was maintained at a constant temperature of 30 °C throughout the entire reaction.

The reaction mixture was pumped by a BT100-2J peristaltic pump (LongerPump Hebei, China) with a flux of 0.5 mL/min through the SpinsolveTM system. The reaction was monitored for approximately 16 h.

2.2.2. Determination of Reaction Rate and Selectivity Factor. The reaction rate was determined by in-line NMR measurements. For this purpose, an enzyme concentration of 1.6 mg/mL was added to the reaction system in a 10 mL reaction mixture. The experiment was carried out at 50, 75, and 100 mmol/mL. The reaction rate of the product formation as well as the selectivity factor α were determined from the concentration data obtained. The measurements were carried out as described in Section 2.2.1.

2.2.3. Complemental Hard Modeling. Since the signals of interest exhibit a high degree of overlap in the measured spectra, a CHM approach is used to effectively determine concentrations of overlapping spectra.[27] For data processing and CHM, the program PEAXACT 5.7 (S-PACT GmbH, Aachen, Germany) was used. A spectral alignment was performed on the DMSO signal (2.8 ppm) in addition to a linear fit baseline subtraction. In addition, a phase correction was performed. PEAXACT was used for deconvolution of the signal since no pure substance of the product LEβPS is commercially available; a CHM is a sufficient way for the deconvolution of these signals.

2.3. Off-Line Measurements. *2.3.1. HPLC Method.* The L-*threo*-β-phenylserine concentration was determined by HPLC on the column Chirex 3126 D-penicillamine (150 × 4.6 mm) (Phenomenex, Aschaffenburg, Germany) with an eluent consisting of 70% 2 mM copper sulfate in water and 30 vol % methanol. The flow rate was 1 mL/min, and detection was carried out at 254 nm at a column temperature of 50 °C. The typical retention time for L-*threo*-β-phenylserine was 16 min.

2.3.2. GC Method. The benzaldehyde concentration was determined by gas chromatography (GC-FID) on the column HP-5 (30 m × 0.250 mm; film thickness is 250 nm) (Agilent Technologies, Inc., Santa Clara, United States) with a flame ionization detector. The carrier gas was helium (1 mL/min, split ratio of 50). Samples were injected at an oven temperature of 135 °C, which was increased to 200 °C at a

https://doi.org/10.1021/acs.oprd.4c00076
Org. Process Res. Dev. 2024, 28, 3791−3800

rate of 10 °C/min. The retention time of benzaldehyde was approximately 6.3 min.

2.3.3. Derivatization and GC−MS Analysis. To confirm the existence of both diastereomers resulting from the threonine aldolase-catalyzed reaction, a sample was taken, derivatized, and analyzed using GC−MS. Derivatization was performed according to the protocol of Urbach et al. using *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA).[29] GC−MS analysis was performed on column DB-5MS (30 m × 0.25 mm; film thickness is 250 nm) (Agilent Technologies, Inc., Santa Clara, United States) with MS. The carrier gas was helium (1 mL/min, splitless). Samples were injected at an oven temperature of 50 °C that was increased to 350 °C at a rate of 5 °C/min. To perform the MS, electrical ionization was utilized at an energy level of 70 electron volts. The acquisition was conducted in positive mode, and the scanning range for the mass-to-charge ratio (m/h) was set between 50 and 800. Based on the GC−MS analysis, the two diastereomeric products were detected at retention times of approximately 25.90 and 26.09 min.

2.3.4. High-Field NMR. ^{1}H NMR measurements were carried out with the 600 MHz NMR system Bruker AVII-600 (Bruker BioSpin GmbH; Billerica, United States) using 16 scans, a sweep of 12019.23 Hz, and an acquisition time of 2.73 s. The samples were diluted in a ratio of 1:1 with D_2O.

3. RESULTS

3.1. Spectral Analysis. The two diastereomers of *β*-phenylserine, namely, L-*threo*-*β*-phenylserine and L-*erythro*-*β*-phenylserine, differ in the absolute configuration at C3. In an NMR spectrum, the differences between these two diastereomers can be observed in their chemical shift values. Furthermore, significant differences between high-field and low-field NMR systems can be observed.

A spectrum obtained from the monitoring of the reaction low-field NMR system NMR Spinsolve 80 MHz ULTRA is displayed below (see Figure 2a). In comparison to that, a spectrum measured by the high-field NMR system Bruker Avance I 600 MHz is shown in Figure 2b. The low-field NRM spectrum was acquired after 5 h of the reaction at a flow rate of 0.3 mL/min. The signal at 9.8 ppm in the NMR spectrum corresponds to the proton of the aldehyde group of benzaldehyde (19), while the signals between 7.9 and 7.4 ppm correspond to the protons of the aromatic group (21, 22, 23, 24, and 25) of benzaldehyde.

The high-field NMR sample was withdrawn after 5 h and prepared for high-field NMR measurements as outlined previously in the Experimental Section and subsequently measured.

In direct comparison, it can be observed that the aromatic region of the L-*threo*-*β*-phenylserine and L-*erythro*-*β*-phenylserine is not in detail resolved at 80 MHz (two singlets) in contrast to a resolved set of peaks at 600 MHz. This is caused by the fact, that the Larmor precession is greater in low-field systems, causing the spectra to be second order. This causes a decrease in information density caused by low magnetic field strength. Although a change from first order to second order results in information loss, both diastereomers can still be identified in the low-field NMR spectra in the aromatic region since there are still two signals detected. These two signals are caused by different chemical environments of protons in the two diastereomers. This can be caused by various reasons.

On the one hand, hydrogen bonding effects could cause a different chemical shift for the diasteromers. The polarity of oxygen and nitrogen can strongly influence the shift of the protons at the *α*- and *β*-carbon. A proton signal at the *β*-carbon (positions 3 and 12) can be detected at approximately 5.15 ppm; however, these signals are relatively weak and show significant overlap compared to the signals of the aromatic group. This is due to the second order nature of the low-field NMR spectra. In the high-field NMR results, the signals are at about 5.275 ppm (12) and about 5.2 ppm (3). Furthermore, in the low-field NMR spectrum at approximately 3.90 and 3.70 ppm, two doublets are observable, corresponding to the protons at the *α*-carbon (positions 2 and 11). These signals arise due to the orientation of the hydroxyl group in the respective diastereomers. The hydroxyl group at position 11 of L-*erythro*-*β*-phenylserine is spatially aligned next to the proton, deprotecting it due to high polarity and causing a shift of its signal upfield. The signal of the proton at position 2 experiences a weaker shift as the distance to the hydroxyl group is greater due to the alignment of the functional groups. This allows L-*erythro*-*β*-phenylserine to be distinguished at position 2. Also, steric effects could be caused by steric hindrance through different special alignments, ultimately causing a chemical shift.[30] For similar reasons, there is also a strong overlap of the signals of the protons in the aromatic groups of the diastereomers. While the high-field NMR spectrum shows a clear separation of the signals, only two strongly overlapping signals can be identified in the low-field spectrum (5, 6, 7, 8, 9, 14, 15, 16, 17, and 18).

The coupling patterns in a NMR spectrum reflect the interactions between adjacent nuclei, and the differences in the coupling patterns between the two diastereomers can reveal the differences in the distances and angles between adjacent nuclei. Nevertheless, since the magnetic field strength is comparatively low, it cannot be determined from the measured second order signals in the low-field NMR spectra.

In the following, the spectral results over time of the in-line NMR measurements are discussed. As glycine is present in excess in the system and benzaldehyde is the limiting component for the reaction, the signals from the protons of the aromatic group of the benzaldehyde are used for further reaction monitoring.

Throughout the reaction period of 16 h, a total of 290 spectra were recorded with an integration time of 2 min per spectrum. There is a decrease in the signals of benzaldehyde (9.8 and 7.9−7.35 ppm), whereas the integral of the aromatic signals of the diastereomeric products shows an increase (7.35−7.25 ppm), as illustrated in the spectra depicted in Figure 2. The aromatic signals of the diastereomers exhibit the strongest band intensity compared to their other proton signals. As a result, these signals are given primary consideration for further experimental evaluations. This approach helps to exclude the influence of background noise in comparison to the protons present at the *α*- and *β*-carbon. To show the influence of the background noise on the signals of the proton at the *α*-carbon, the concentration data from the areas of the proton at the *α*-carbon can be found in the Supporting Information. However, since a strong overlap of the signals of the diastereomeric products can be recognized, CHM is used to perform data processing as described in the Experimental Section part.

3.2. Concentration Determination by CHM. The NMR data was used to determine concentrations, with signals from

https://doi.org/10.1021/acs.oprd.4c00076
Org. Process Res. Dev. 2024, 28, 3791−3800

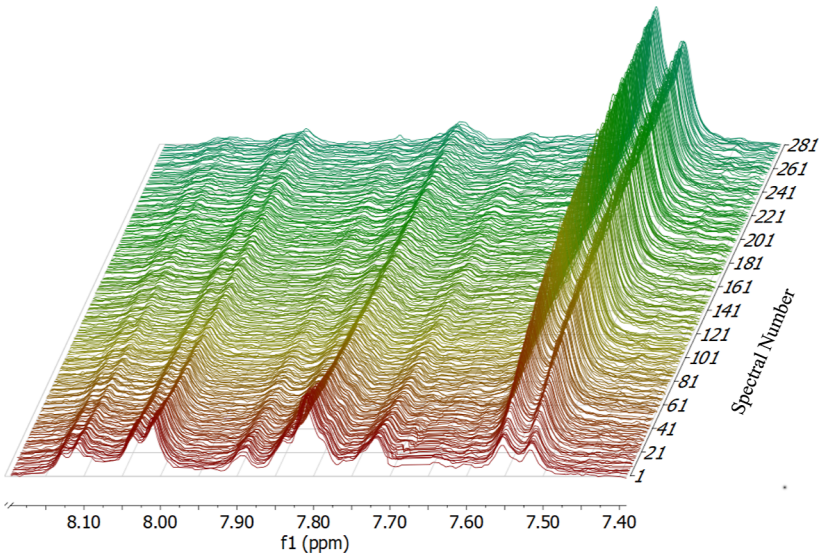


Figure 3. Reaction monitoring of the aldol reaction measured by Magritek Spinsolve 80 ULTRA (Magritek, Aachen; Germany), all ^{1}H spectra were measured at 80 MHz.

the aromatic region between 7.90 and 7.40 ppm for benzaldehyde and 7.35 and 7.20 ppm for the two diastereomeric products. The signals present in the aromatic region show a higher signal intensity at lower concentrations in comparison to the signals of the protons at positions 2 and 11 (see Figure 2). This area is subsequently used to determine the concentrations of the diastereomers. Peak deconvolution is carried out by CHM. By observing the more pronounced signals and applying CHM, it should be possible to achieve a higher accuracy of the determined data compared to the signals at positions 2 and 11. Figure 4 shows the concentrations determined from the signals in the in-line NMR measurements by using a CHM approach. The presented results exclude concentration data derived from the signal of the aldehyde group of benzaldehyde. These data can be found in the Supporting Information. In the thermodynamic equilibrium, a concentration of 24.15 mM was determined utilizing the aldehyde group signal, whereas concentrations of 23.61 mM were ascertainable through signals from the aromatic group and the CHM approach. A discrepancy of 2.29% exists between these concentration values. This marginal deviation serves to affirm the precision of the applied CHM approach.

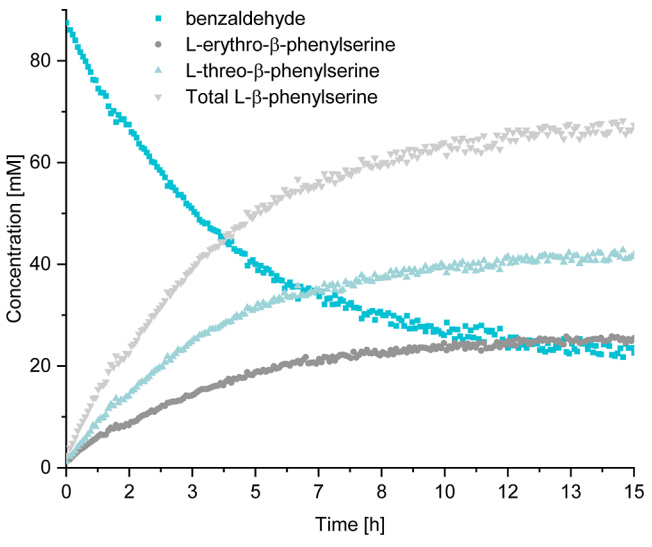


Figure 4. Concentration of L-*erythro*-*β*-phenylserine and L-*threo*-*β*-phenylserine during aldol reaction (1 M glycine; 90 mM benzaldehyde in 20 v/v DMSO in KPi buffer pH 8.0 V = 10 mL; T = 30 °C; t = 15 h), IHM (Peaxact 5.7 S.Pact; Aachen Germany) from spectral data of low-field NMR Spinsolve 80 MHz ULTRA (Magritek, Aachen).

Over time, a decrease in the benzaldehyde concentration is observed, whereas the total L-*β*-phenylserine concentration increases. Additionally, there is a clear increase in the corresponding diastereomer concentration. To provide a more accurate indication of the formation rate of the products, the linear range of the reaction course below 5% conversion is depicted in Figure 8.

The experiments were conducted at varying benzaldehyde concentrations (50, 75, and 90 mM), and the initial reaction rates of the diastereomeric products were determined and compared using the first data points obtained. Figure 5 illustrates the initial rates and selectivities of both diastereomers synthesized in the reaction system, which were determined during the experiments.

As the substrate concentration increases, the overall rate of product formation also increases. However, it is evident that the production of L-*threo*-*β*-phenylserine has increased more significantly than that of L-*erythro*-*β*-phenylserine. This trend is also reflected in the selectivity factors. The selectivity factor refers to the preference or specificity of the enzyme for one diastereomer over the other. It can be defined as the ratio of the reaction rates from the substrate to one specific diastereomer. A selectivity factor greater than 1 indicates a preference for one diastereomeric product, while a selectivity factor less than 1 suggests a preference for the other respective product. A selectivity factor of 1 implies equal production of both diastereomers. The selectivity factors at different benzaldehyde concentrations can be seen in Figure 6. It can be observed that an increase in substrate concentration leads to a higher diastereoselectivity for L-*threo*-*β*-phenylserine and a lower diastereoselectivity for L-*erythro*-*β*-phenylserine. Overall, a change in the diastereoselectivity of threonine aldolases is

https://doi.org/10.1021/acs.oprd.4c00076
Org. Process Res. Dev. 2024, 28, 3791−3800

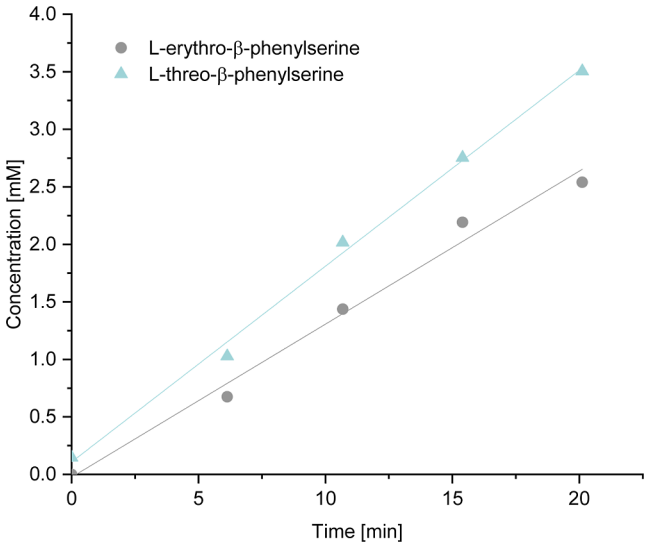


Figure 5. Concentration of L-*erythro*-*β*-phenylserine and L-*threo*-*β*-phenylserine during aldol reaction (reaction conditions identical to Figure 4) and (right) IHM (Peaxact 5.7 S.Pact; Aachen Germany) from spectral data of low-field NMR Spinsolve 80 MHz ULTRA (Magritek, Aachen).

caused by varied substrate concentrations. However, the selectivity factor for L-*threo*-*β*-phenylserine consistently exceeds 1, indicating a stronger preference for the formation of this product compared to L-*erythro*-*β*-phenylserine. In each experiment conducted, the selectivity factor for L-*erythro*-*β*-phenylserine remains below 1.

Besides the focus on reaction rates, attaining a thermodynamically controlled reaction system is also noteworthy. The reaction mixture utilized contained 90 mM of benzaldehyde. Figure 10 illustrates the relationship between the analytical yield of each diastereomer and the diastereoselectivity of the synthesized products in order to gain a deeper understanding of their distribution over time. The diastereoselectivity is calculated as follows

$$S_1 = \frac{c_{D1}}{c_{D1} + c_{D2}} \quad (1)$$

S_1 is the diastereoselectivity toward diastereomer 1, c_{D1} is the concentration of diastereomer 1, and c_{D2} is the concentration of diastereomer 2. The results demonstrate an increase in the diastereoselectivity of L-*threo*-*β*-phenylserine, accompanied by a decrease in the diastereoselectivity of L-*erythro*-*β*-phenylserine. The diastereoselectivity remains constant once the yield reaches approximately 40%, which suggests that the reaction has reached thermodynamic equilibrium. A diastereomeric excess of 37:63 (*erythro*/*threo*) can be observed in the system under thermodynamic control. In addition, the initial validation results can be detected. Off-line results can also be recognized in addition to the benchtop NMR results. The L-*threo*-*β*-phenylserine concentration was determined by HPLC and the benzaldehyde concentration by GC. The concentration of L-*erythro*-*β*-phenylserine was determined from a mass balance. The diastereoselectivities determined from these data are listed in Figure 7. Comparable results can be seen with a diastereomeric excess of 36:64 (*erythro*/*threo*).

Besides the analytical yield, the conversion rate of benzaldehyde is a crucial parameter for designing a process. Figure 8 depicts the connection between the conversion of

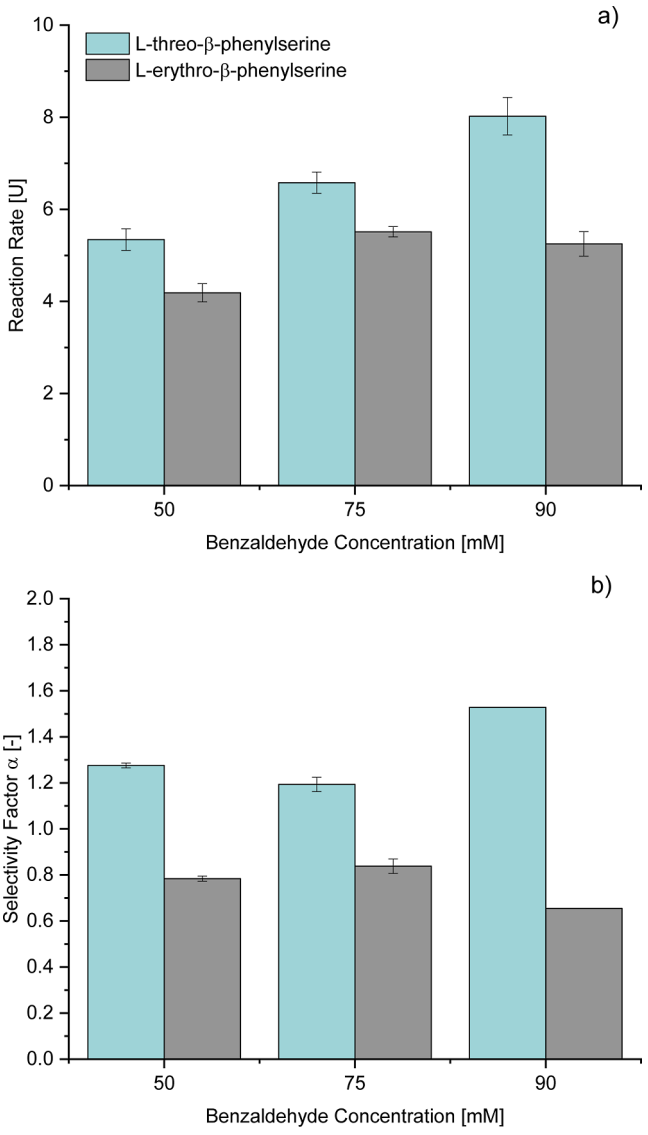


Figure 6. Reaction rates (a) and selectivity factors; (b) of L-*threo*-*β*-phenylserine and L-*erythro*-*β*-phenylserine at initial rates as a function of benzaldehyde concentration.

benzaldehyde and the diastereoselectivity of L-*erythro*-*β*-phenylserine. The analytical yield is defined as follows

$$\text{yield} = \left(\frac{c_{Pn}}{c_{S0}}\right) \times 100 \quad (2)$$

c_{Pn} is the product concentration a at the respective time n and c_{S0} is the substrate concentration at the start of the reaction.

The data reveal a consistent decline in the diastereoselectivity of L-*erythro*-*β*-phenylserine until the substrate conversion reaches 30%. Beyond this point, the diastereoselectivity remains constant in the range of measuring errors of ∼0.002, indicating that thermodynamic equilibrium has been reached.

3.3. Data Validation by GC, HPLC, High-Field NMR, and GC−MS. To verify the accuracy of the NMR data, a combination of GC, HPLC, and high-field NMR techniques were employed. In batch experiments, samples were extracted every 15 min over a 2 h period, and the levels of benzaldehyde and L-*threo*-*β*-phenylserine were evaluated using HPLC and GC. Furthermore, benzaldehyde and phenylserine concen-

https://doi.org/10.1021/acs.oprd.4c00076
Org. Process Res. Dev. 2024, 28, 3791−3800

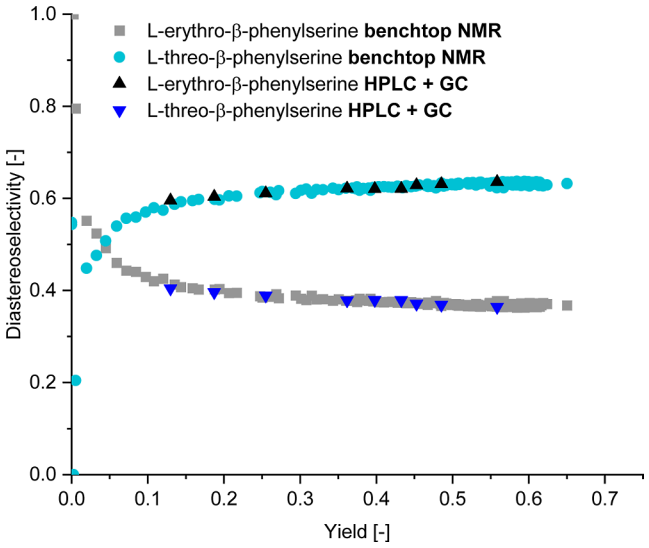


Figure 7. Diastereoselectivity in the formation of L-*erythro*-*β*-phenylserine and L-*threo*-*β*-phenylserine catalyzed by L-threonine aldolase as a function of analytical yield of the respective diastereomeric molecule; NMR Spinsolve 80 MHz ULTRA (Magritek, Aachen).

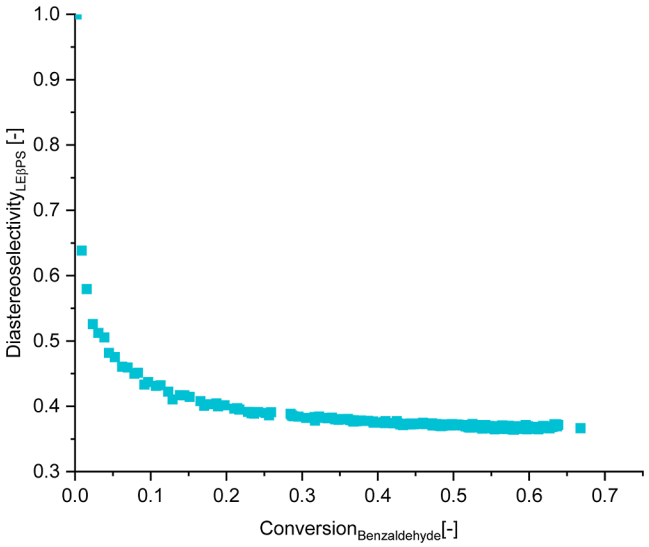


Figure 8. Diastereoselectivity of L-*erythro*-*β*-phenylserine formation as a function of benzaldehyde conversion; spectral data measured by low-field NMR Spinsolve 80 MHz ULTRA (Magritek, Aachen).

trations were determined by off-line high-field NMR measurements. The results of these experiments can be seen in Figure 9.

Phenylserine concentration was determined based on HPLC and GC data. Since calibration for L-*erythro*-*β*-phenylserine was unattainable due to the unavailability of the pure substance commercially, benzaldehyde concentration was determined through GC, while L-*threo*-*β*-phenylserine concentration was determined through HPLC. The L-*erythro*-*β*-phenylserine concentration of the total phenylserine content was calculated through mass balance.

Despite the likelihood of a higher degree of inaccuracy due to errors introduced by extraction, GC measurement, and HPLC measurement, a strong correlation of the data for both phenylserine and benzaldehyde determinations is evident. This

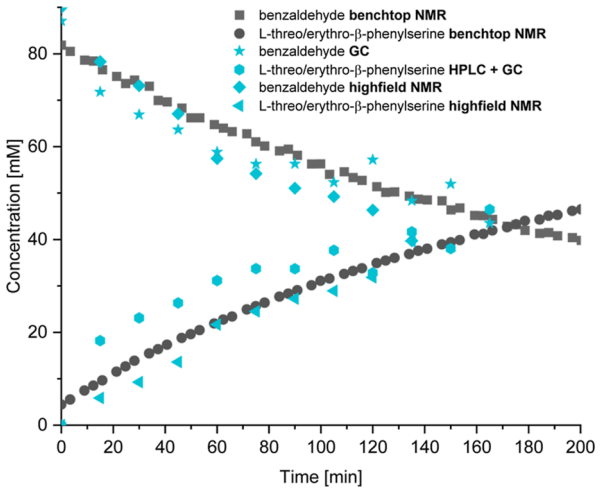


Figure 9. Validation of NMR by comparison to GC and HPLC data in the case of a batch synthesis; spectral data measured by low-field NMR Spinsolve 80 MHz ULTRA (Magritek, Aachen).

indicates the robustness of the benchtop NMR analysis for the chosen reaction system.

It is worth noting that the disparity in data between high- and low-field NMR is less than that between low-field NMR and the combined HPLC and GC methods. This observation further implies that the HPLC and GC methods may be subject to more significant measurement-related errors. Moreover, an observable displacement of approximately 5 mM is observed in the benchtop NMR data for phenylserine. This displacement can be attributed to the distinctive noise generated during the online measurements. Additionally, there is a minor overlap between the aromatic benzaldehyde and aromatic phenylserine functional groups. When employing the CHM, the concentration of the product is determined based on the substrate signals, which causes a determination of product concentration at the beginning of the reaction. Subsequently, in order to provide further substantiation of the existence of diastereomers, a sample from the experiment underwent derivatization with MSTFA and was subjected to GC−MS analysis, as outlined in the Experimental Section. Figure 10 displays the measured signals and mass spectra, which reveal two distinct signals with retention times of approximately 25.90 and 26.09 min. The fragmentation patterns of the diastereomeric phenylserine products closely align with the measured mass spectra, thereby indicating the presence of L-*threo*-*β*-phenylserine and L-*erythro*-*β*-phenylserine. Detailed fragmentation patterns of these two diastereomeric amino acids can be found in the Supporting Information.

4. DISCUSSION

As a result of the low magnetic field in comparison to that of high-field NMR, complete signal separation of the diastereomeric products could not be achieved, as depicted in Figure 2. This is due to the second order nature of the weak-field NMR spectra. In direct comparison, it was possible to separate the signals in the 600 MHz high-field NMR spectrum (Figure 3). Nevertheless, by using CHM, it was possible to identify the "hidden" regions for the determination of the concentration. By this, an underestimation of the concentration values could be prevented, which would have introduced a notable source of systematic error into the analysis. This CHM approach was

https://doi.org/10.1021/acs.oprd.4c00076
Org. Process Res. Dev. 2024, 28, 3791−3800

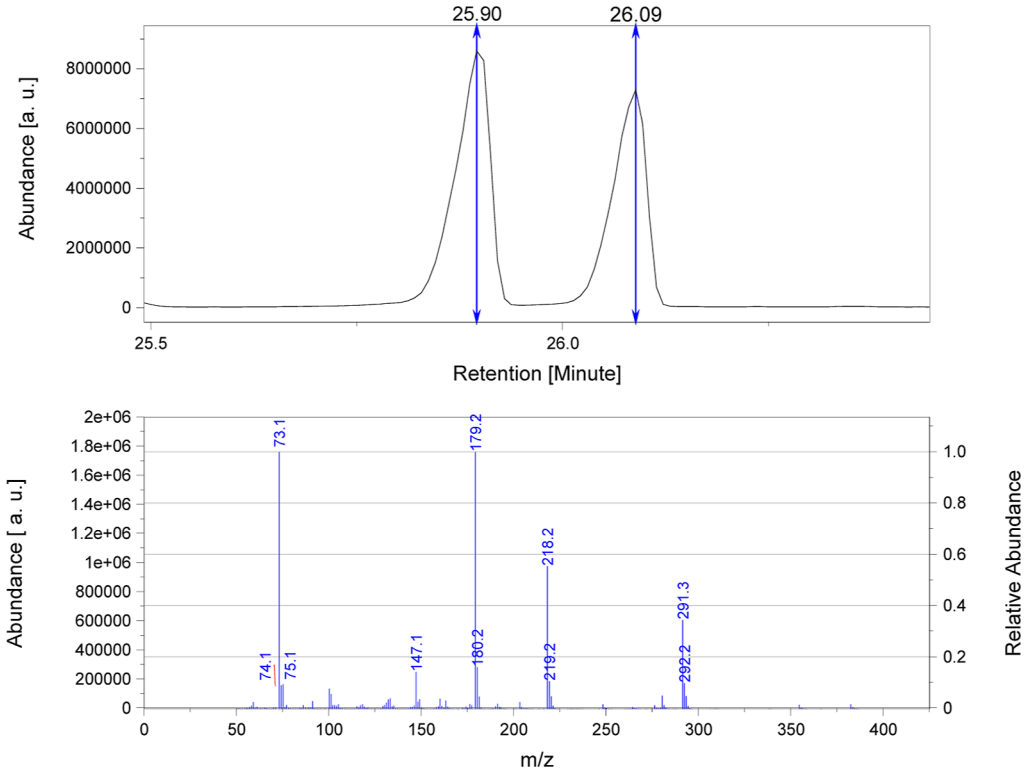


Figure 10. GC−MS data of reaction mixture taken after 6 h subsequent to derivatization with MSTFA. Both detected compounds have the same mass spectrum.

also applied to the signals of the aromatic region of benzaldehyde; compared to the concentration of the signal from the proton on the aldehyde, there was a deviation of 2.29%, indicating a robust CHM validation experiments were conducted to confirm this, whereby samples were measured with a 600 MHz high-field NMR.

Additionally, off-line measurements were performed using HPLC to determine L-*threo*-*β*-phenylserine concentrations and GC measurements to determine benzaldehyde concentrations, as shown in Figure 9. The obtained concentration values align with the benchtop NMR results. However, a discrepancy can be observed between the phenylserine concentration values obtained from the high-field NMR system and those from the benchtop NMR system. Nevertheless, these data are within an acceptable range. One reason for this deviation could be a loss of information due to the second order of the low-field NMR spectra; although a CHM was applied, a loss may have occurred due to the shifting of several signals. In spite of this, there is only a small deviation between the low-field and high-field NMR. The concentrations determined through HPLC and GC exhibit a greater degree of error, most likely due to measurement inaccuracies and the extraction process during sample preparation used for the determination of benzaldehyde. Nonetheless, the obtained data remain within a range comparable to that of the benchtop NMR data. Incorporating the high-field NMR data further supports the assumption of higher measurement errors in the HPLC and GC measurements. The benchtop NMR data and high-field NMR data show a stronger correlation with each other. Although the signals of the diastereomers were not completely separated in the low-field NMR measurements, they were still detectable. The use of a CHM allowed for the determination of concentrations from the signals of the aromatic region of benzaldehyde and its two diastereomeric products, as shown in Figures 3 and 6. Despite signal overlap, the concentrations could be determined, enabling in-line monitoring of the synthesis of these diastereomeric products.

The concentration curves allowed for the measurement of the initial reaction rates of product formation at various benzaldehyde concentrations. In general, it was observed that as substrate concentration increased, the diastereoselectivity of L-*threo*-*β*-phenylserine increased while that of LE*β*PS decreased. This phenomenon could be attributed to several factors. First, a higher substrate concentration can lead to an increased reaction rate. At high substrate concentrations, the synthesis of L-*erythro*-*β*-phenylserine may proceed too rapidly for accurate NMR measurements because of the integration time needed. A substrate surplus inhibition of the active site for the synthesis of L-*erythro*-*β*-phenylserine is a plausible explanation, which may favor higher L-*threo*-*β*-phenylserine product formation rates. Overall substrate inhibition by benzaldehyde is already proven in the literature.[11] The study further investigated the change in the diastereoselectivity of the reaction system. Figure 7 illustrates the diastereoselectivity of the two diastereomeric products at an initial substrate concentration of 90 mM. It is evident that the diastereoselectivity for L-*threo*-*β*-phenylserine increases while the diastereoselectivity for the LE*β*PS decreases. This was also validated by off-line GC and HPLC data. In thermodynamic equilibrium, a diastereomeric excess of 37:63 (*erythro*/*threo*) was determined by in-line NMR measurements and 36:64 (*erythro*/*threo*) by HPLC and GC (see Figure 7). These datasets coincide very closely. Once the analytical yield reaches approximately 30%, the diastereoselectivities remain constant, indicating that a thermodynamically controlled reaction system has been established.

https://doi.org/10.1021/acs.oprd.4c00076
Org. Process Res. Dev. 2024, 28, 3791−3800

Similar ratios of 63:37 (*erythro*/*threo*) have also been reported in the literature at a maximum benzaldehyde conversion of 80%.[13,14] Figure 8 illustrates that the diastereoselectivity of L-*erythro*-β-phenylserinedecreases as the conversion of benzaldehyde increases. It can be observed that a thermodynamically controlled system is already established at around 30% conversion, as indicated by the constant diastereoselectivity.

5. CONCLUSIONS

The enzymatic synthesis of L-*threo*-β-phenylserine and L-*erythro*-β-phenylserine was successfully determined by in-line NMR for the first time. Overlapping signals by a weaker electromagnetic field and high background noise due to in-line measurements represent a challenge. However, a CHM approach was employed to overcome this issue, and while an improvement in the background noise would be desirable, the current quality of the spectra obtained with benchtop NMR instruments is sufficient for process development. The combination of stereoselective biocatalysis and in-line NMR technologies has great potential for the rapid optimization of enzymatic syntheses to enable the sustainable production of pharmaceutical components.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.oprd.4c00076.

Gene sequence of the plasmid employed for the synthesis of L-threonine aldolase; amino acid sequence of L-threonine aldolase; mass spectrum and fragmentation patterns of the derivatized phenylserine; concentration profiles for LEβPS and L-*threo*-β-phenylserine; concentration data derived from the proton signal of the aldehyde group of benzaldehyde exhibit comparable characteristics to those derived from the aromatic group; and COSY measurements using high- and low-field NMR (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Andreas Liese − *Institute of Technical Biocatalysis, Hamburg University of Technology, 21073 Hamburg, Germany;* orcid.org/0000-0002-4867-9935; Email: liese@tuhh.de

Authors

Luca F. Schmidt − *Institute of Technical Biocatalysis, Hamburg University of Technology, 21073 Hamburg, Germany;* orcid.org/0000-0002-2203-096X

Logia Jolly − *Chair of Industrial Organic Chemistry and Biotechnology, University of Bielefeld, 33615 Bielefeld, Germany*

Leon Hennecke − *Institute of Technical Biocatalysis, Hamburg University of Technology, 21073 Hamburg, Germany*

Fernando Lopez Haro − *Institute of Technical Biocatalysis, Hamburg University of Technology, 21073 Hamburg, Germany*

Harald Gröger − *Chair of Industrial Organic Chemistry and Biotechnology, University of Bielefeld, 33615 Bielefeld, Germany;* orcid.org/0000-0001-8582-2107

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.oprd.4c00076

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are very grateful to the "Deutsche Forschungsgemeinschaft" (DFG) for their financial support (grant no. GR 3461/10-1). We also gratefully acknowledge the contributions of Leon Klose, Grit Brauckmann, Kassim Patschtun, and Thorsten Mix. Leon Klose provided valuable assistance in the GC-FID and GC−MS methods, while Grit Brauckmann and Kassim Patschtun, as bachelor students and BTA interns, respectively, greatly supported our laboratory work. Their involvement was essential to the successful completion of this study. Additionally, thank Thorsten Mix from the University of Hamburg for his valuable assistance in applying high-field NMR measurement.

■ REFERENCES

(1) Waldeck, B. Biological Significance of the Enantiomeric Purity of Drugs. *Chirality* **1993**, *5*, 350−355.

(2) Gijsen, H. J. M.; Qiao, L.; Fitz, W.; Wong, C.-H. Recent Advances in the Chemoenzymatic Synthesis of Carbohydrates and Carbohydrate Mimetics. *Chem. Rev.* **1996**, *96*, 443−474.

(3) Sheldon, R. A.; Woodley, J. M. Role of Biocatalysis in Sustainable Chemistry. *Chem. Rev.* **2018**, *118*, 801−838.

(4) Dückers, N.; Baer, K.; Simon, S.; Gröger, H.; Hummel, W. Threonine aldolases—screening, properties and applications in the synthesis of non-proteinogenic β-hydroxy-α-amino acids. *Appl. Microbiol. Biotechnol.* **2010**, *88*, 409−424.

(5) Fesko, K. Threonine Aldolases: Perspectives in Engineering and Screening the Enzymes with Enhanced Substrate and Stereo Specificities. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 2579−2590.

(6) Gwon, H.-J.; Baik, S.-H. Diastereoselective Synthesis of L-threo-3,4-Dihydroxyphenylserine by Low-Specific L-threonine Aldolase Mutants. *Biotechnol. Lett.* **2010**, *32*, 143−149.

(7) Liu, J. Q.; Odani, M.; Yasuoka, T.; Dairi, T.; Itoh, N.; Kataoka, M.; Shimizu, S.; Yamada, H. Gene Cloning and Overproduction of Low-Specificity D-Threonine Aldolase from Alcaligenes xylosoxidans and Its Application for Production of a Key Intermediate for Parkinsonism Drug. *Appl. Microbiol. Biotechnol.* **2000**, *54*, 44−51.

(8) Rocha, J. F.; Pina, A. F.; Sousa, S. F.; Cerqueira, N. M. F. S. A. PLP-Dependent Enzymes as Important Biocatalysts for the Pharmaceutical, Chemical, and Food Industries: A Structural and Mechanistic Perspective. *Catal. Sci. Technol.* **2019**, *9*, 4864−4876.

(9) Steinreiber, J.; Fesko, K.; Reisinger, C.; Schürmann, M.; van Assema, F.; Wolberg, M.; Mink, D.; Griengl, H. Threonine Aldolases—An Emerging Tool for Organic Synthesis. *Tetrahedron* **2007**, *63*, 918−926.

(10) Machajewski, T. D.; Wong, C. The Catalytic Asymmetric Aldol Reaction. *Angew. Chem., Int. Ed.* **2000**, *39*, 1352−1375.

(11) Baer, K. *Die Stereoselektive Aldolreaktion in Biotransformationen und Chemoenzymatischen Eintopfsynthesen*. Dissertation, Friedrich-Alexander-Universitaet, 2011.

(12) Shibata, K.; Shingu, K.; Vassiley, V. P.; Nishide, K.; Fujita, T.; Node, M.; Kajimoto, T.; Wong, C. Kinetic and Thermodynamic Control of L-Threonine Aldolase Catalyzed Reaction and Its Application to the Synthesis of Mycestericin D. *Tetrahedron Lett.* **1996**, *37*, 2791−2794.

(13) Gutierrez, M. L.; Garrabou, X.; Agosta, E.; Servi, S.; Parella, T.; Joglar, J.; Clapés, P. Serine Hydroxymethyl Transferase from *Streptococcus thermophilus* and L-Threonine Aldolase from *Escherichia coli* as Stereocomplementary Biocatalysts for the Synthesis of β-Hydroxy-α,ω-diamino Acid Derivatives. *Chem.—Eur. J.* **2008**, *14*, 4647−4656.

(14) Kimura, T.; Vassilev, V. P.; Shen, G.-J.; Wong, C.-H. Enzymatic Synthesis of β-Hydroxy-α-amino Acids Based on Recombinant d- and l-Threonine Aldolases. *J. Am. Chem. Soc.* **1997**, *119*, 11734−11742.
(15) Wozniak, T. J.; Bopp, R. J.; Jensen, E. C. Chiral Drugs: An Industrial Analytical Perspective. *J. Pharm. Biomed. Anal.* **1991**, *9*, 363−382.
(16) *Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*, 2nd ed.; Bakeev, K. A., Ed.; Wiley: Chichester, 2010.
(17) *Specification of Drug Substances and Products: Development and Validation of Analytical Methods*, 2nd ed.; Riley, C. M., Rosanske, T. W., Reid, G. L., Eds.; Elsevier: Amsterdam, 2020.
(18) Castaing-Cordier, T.; Bouillaud, D.; Farjon, J.; Giraudeau, P. Recent Advances in Benchtop NMR Spectroscopy and Its Applications. *Annu. Rep. NMR Spectrosc.* **2021**, *103*, 191−258.
(19) Reich, H. Hans Reich's Collection. NMR Spectroscopy. https://organicchemistrydata.org/hansreich/resources/nmr/?page=nmr-content%2F (accessed Nov 13, 2023).
(20) Blümich, B. Low-Field and Benchtop NMR. *J. Magn. Reson.* **2019**, *306*, 27−35.
(21) Castaing-Cordier, T.; Ladroue, V.; Besacier, F.; Bulete, A.; Jacquemin, D.; Giraudeau, P.; Farjon, J. High-Field and Benchtop NMR Spectroscopy for the Characterization of New Psychoactive Substances. *Forensic Sci. Int.* **2021**, *321*, 110718.
(22) Grootveld, M.; Percival, B.; Gibson, M.; Osman, Y.; Edgar, M.; Molinari, M.; Mather, M. L.; Casanova, F.; Wilson, P. B. Progress in Low-Field Benchtop NMR Spectroscopy in Chemical and Biochemical Analysis. *Anal. Chim. Acta* **2019**, *1067*, 11−30.
(23) Rudszuck, T.; Nirschl, H.; Guthausen, G. Perspectives in Process Analytics Using Low Field NMR. *J. Magn. Reson.* **2021**, *323*, 106897.
(24) Silva Elipe, M. V.; Milburn, R. R. Monitoring Chemical Reactions by Low-Field Benchtop NMR at 45 MHz: Pros and Cons. *Magn. Reson. Chem.* **2016**, *54*, 437−443.
(25) Spiegelman, C. H.; Bennett, J. F.; Vannucci, M.; McShane, M. J.; Coté, G. L. A Transparent Tool for Seemingly Difficult Calibrations: The Parallel Calibration Method. *Anal. Chem.* **2000**, *72*, 135−140.
(26) Kriesten, E.; Alsmeyer, F.; Bardow, A.; Marquardt, W. Fully Automated Indirect Hard Modeling of Mixture Spectra. *Chemom. Intell. Lab. Syst.* **2008**, *91*, 181−193.
(27) Kriesten, E.; Mayer, D.; Alsmeyer, F.; Minnich, C. B.; Greiner, L.; Marquardt, W. Identification of Unknown Pure Component Spectra by Indirect Hard Modeling. *Chemom. Intell. Lab. Syst.* **2008**, *93*, 108−119.
(28) Liu, J. Q.; Dairi, T.; Itoh, N.; Kataoka, M.; Shimizu, S.; Yamada, H. Gene Cloning, Biochemical Characterization and Physiological Role of a Thermostable Low-Specificity L-Threonine Aldolase from Escherichia coli. *Eur. J. Biochem.* **1998**, *255*, 220−226.
(29) Urbach, D. MSTFA/MSTFA-d9 Derivatization of Amphetamine for GC/MS Detection and Identification. https://www.sigmaaldrich.com/DE/de/technical-documents/protocol/analytical-chemistry/gas-chromatography/mstfa-d9-derivatization (accessed Feb 23, 2023).
(30) Günther, H. *NMR Spectroscopy: Basic Principles, Concepts, and Applications in Chemistry*, 2nd ed.; Wiley: Chichester, 2001.

https://doi.org/10.1021/acs.oprd.4c00076
Org. Process Res. Dev. 2024, 28, 3791−3800