Determining Inversion Barriers in Atropisomers – A Tutorial for Organic Chemists

Michel Rickhaus, Lukas Jundt, and Marcel Mayor*

Abstract: Dynamic behavior is a fascinating property of natural and artificial systems and its understanding has significantly impacted the transformation of molecular interchanges into controlled molecular motion. In this tutorial, the key descriptors of enantiomeric stability are examined in-depth. Enantiomerization and racemization are discussed and differentiated on a fundamental level proposing a unified and distinct nomenclature. Their mathematical meanings and relations are described and deduced cohesively in the context of atropisomerization. The calculation of inversion barriers from thermodynamic and kinetic data is demonstrated and the interdependences between the latter are explained mathematically. Using current examples from our group, the determination of rate constants and the thermodynamic parameters is shown in a step-by-step manner using the most common techniques. The tutorial is concluded with aspects and considerations concerning statistical data analysis and error determination of measurements including a practical guide to Monte-Carlo simulations.

Keywords: Enantiomerization · Inversion barrier · Kinetics · Racemization · Tutorial

1. Introduction

Dynamic behavior is a key property in the nano-engineered world. Be it molecular motors, switches, or capsules, all of them rely on interchanging states, leading to features like propulsion, controlled motion and uptake/release cycles of targets. Understanding the dynamics that govern these systems allows the design of more capable systems. But it is not only in the design of new molecular machines that dynamic behavior is important. We find it in pharmacology and biology as readily, be it protein-folding, enzyme kinetics or cooperativity. Consequently a vast number of rationalizations have emerged, each with their own background and terminology. Chemists have the unique opportunity to study this behavior on a fundamental level. From single bond rotations to foldamers, model systems showing a finite level. The reaction coordinate of such systems typically includes a barrier that is a measure of how readily one conformer converts into its mirror image. This barrier is a key descriptor for the stability of handedness and the design of cascading interchanges that allow, for instance, unimolecular motion.

In the shape of a tutorial we present herein a collection of formulas, methods and procedures to efficiently investigate the thermodynamic parameters of the dynamic interchange of enantiomers, particularly the barrier of inversion. With its uniform nomenclature and illustrated examples, it is intended to serve in particular organic chemists as an easy-to-follow reference.

2. Mathematical Toolkit

2.1 Enantiomerization, Racemization and the Rate Constant

Mathematical toolsets addressing the kinetics of a racemization process exist, and describe racemization usually as a 1st order process. However, the interested reader is soon lost in inconsistent terminology and imprecisely defined formulas and derivatives thereof. While the information is undoubtedly there and many excellent books highlight the various aspects of dynamic interconversion in detail, we reasoned that it would be convenient to collect the most important ideas, formulas and transformations in a mathematical toolkit with a consistent terminology.

The most important distinction to make is to treat the process of losing chiral information, that is turning an enantiomerically pure sample into a racemate (racemization) and the interconversion of one enantiomer into the other (enantiomerization) separately. These two processes are closely related as will be outlined, but are often incorrectly used as synonyms.

Enantiomerization is a microscopic process describing the reversible conversion of one enantiomer (A) into the other (B) on a microscopic level.

\[
A \xrightarrow{k_{A\to B}} B \quad k_{B\to A}
\] (1)

Racemization on the other hand is a macroscopic process and describes the conversion of a collection of enantiomers (A) into a racemic mixture (A+B) with an A:B ratio of 1:1.

\[
A \xrightarrow{k_{rac}} [50\% A + 50\% B] \xleftarrow{k_{rac}} B
\] (2)

The process of racemization is prominent in literature but is technically a special case of the more general enantiomerization with some boundary conditions: The rate constant for the forward process \(k_{A\to B}\) is equal to the backward process \(k_{B\to A}\), and the enantiomers A and B show a 1:1 ratio in the equilibrium. Enantiomerization can lead to racemates but does not necessarily
need to do so. Important in the context of atropisomerization is that $k_{rac}$ is deduced experimentally since it corresponds to a macroscopic system and is often observed as loss of a chiral signal but $k_e$ (the microscopic system) should be used when determining the thermodynamic parameters of activation, that is, $ΔG^0_{rac}$, $ΔH^0_{rac}$, and $ΔS^0_{rac}$. Fortunately, enantiomerization and racemization can be easily interconverted.[7] For an achiral environment which can reach a racemate:

$$k_{A\rightarrow B} = k_{B\rightarrow A} = k_e$$  (3)

which ultimately results in

$$2k_e = k_{rac}$$  (4)

since the interconversion of one molecule reduces the enantiomeric excess by two molecules.

Over the course of this tutorial we will only consider processes of systems that a) do not describe diastereomerization; b) show equal rates of the forward and the back reaction that can be described as the rate of enantiomerization ($k_{A\rightarrow B} = k_{B\rightarrow A}$) and are therefore capable of reaching a racemic state; c) are taking part in a dynamic exchange of enantiomers (or vice versa). The rate law shows the exponential decay of a pure enantiomer:

$$[A]_t = [A]_0 e^{-2k_e t}$$  (17)

The half life of a pure enantiomer is given by:

$$t_{1/2} = \frac{ln 2}{2k_e} \approx 0.693 \frac{1}{2k_e}$$  (20)

The half live is a valuable parameter as it directly describes the stability of an enantiomer over time.

### 2.4 The Inversion Barrier

A – if not the – key value to indicate the stability of enantiomers is the inversion barrier ($ΔG^‡$) which describes the Gibbs free energy of activation usually denoted $ΔG^0_e$ or (imprecisely) $ΔG^†_{rac}$. The best practice is to use $ΔG^0_e$ as the dagger ($‡$) indicates that the value describes the free energy change of activation (Fig. 1a), *i.e.* going from A to the transition state (A → A$^\ddagger$) while the $e$ specifies that the process under investigation is converting one particle of A over the transition state A$^\ddagger$ to one particle of B with inversed configuration (enantiomerization).

It can be stated that $ΔG^‡_{e(T)}$ is dependent on temperature and consists of an enthalpic and an entropic term:

$$ΔG^‡_{e(T)} = ΔH^‡_{e} - T ΔS^‡_{e}$$  (21)

The rate constant $k_e$ can be related to $ΔG^‡_{e(T)}$ by the Eyring equation:

$$k_e = \frac{k_B T}{h} e^{-\frac{ΔG^‡_{e(T)}}{RT}}$$  (22)
Solved for $\Delta G_e^{\dagger}$:

$$\Delta G_e^{\dagger}(T) = RT \left[ \ln \left( \frac{k_g h}{k_e} \right) - \ln (k_e) \right]$$  \hspace{1cm} (23)

Instead of using the Gibbs free energy the activation energy, $E_a$, is often found especially in older literature. The activation energy can be related to $k_e$ by the Arrhenius equation:

$$k_e = A e^{-\frac{E_a}{RT}}$$  \hspace{1cm} (24)

$E_a$ and $\Delta G_e^{\dagger}$ can now be related by:

$$E_a = \Delta H_e^{\dagger} + RT$$  \hspace{1cm} (25)

Substitution of Eqn. (25) into Eqn. (21) gives

$$\Delta G_e^{\dagger}(T) = E_a - T \left( \Delta H_e^{\dagger} + \Delta S_e^{\dagger} \right)$$  \hspace{1cm} (26)

2.5 Enthalpy and Entropy of Enantiomerization

To determine values for $\Delta H_e^{\dagger}$ and $\Delta S_e^{\dagger}$ one can use the temperature dependence of $k_e$. Substituting $\Delta G_e^{\dagger}$ for Eqn. (21) gives the expanded Eyring equation:

$$k_e = \left( \frac{k_g h}{k_e} \right) e^{-\frac{\Delta H_e^{\dagger} + \Delta S_e^{\dagger}}{RT}}$$  \hspace{1cm} (27)

which can be brought into a linear form:

$$\ln \left( \frac{k_e}{T} \right) = -\frac{\Delta H_e^{\dagger}}{R} \left( \frac{1}{T} \right) + \frac{\Delta S_e^{\dagger}}{R} + \ln \left( \frac{k_g h}{k_e} \right)$$  \hspace{1cm} (28)

Although the formula appears convoluted, it essentially states that plotting $ln(k_e/T)$ vs. $1/T$ will yield a straight line with slope $a = -(\Delta H_e^{\dagger}/R)$ and the intersection $b = \Delta S_e^{\dagger}/R + \ln(k_g/h)$ (Eyring plot). Solved for the corresponding thermodynamic parameters:

$$\Delta H_e^{\dagger} = -aR$$  \hspace{1cm} (29)

$$\Delta S_e^{\dagger} = R \left[ b - \ln \left( \frac{k_g h}{k_e} \right) \right] = -R \left[ b - 23.8 \right]$$  \hspace{1cm} (30)

from which $\Delta G_e^{\dagger}$ can be calculated for a specific temperature (Eqn. (21)). Comparing the barrier directly obtained from the rate constant (using Eqn. (22)) at a specific temperature ($\Delta G_e^{\dagger}(T)$) to the one from the Eyring regression (Eqn. (21)) using the same temperature is good practice to spot errors in the Eyring regression.

If both $E_a$ and $\Delta G_e^{\dagger}$ are known, the entropy can also be determined by using Eqn. (26) and solving for $\Delta S_e^{\dagger}$:

$$\Delta S_e^{\dagger} = \frac{E_a - \Delta G_e^{\dagger}(T)}{T}$$  \hspace{1cm} (31)

3. Experimental Determination of Rate Constants – Three Common Options

When tasked with the determination of the thermodynamic parameters of enantiomerization, or more generally, with the optical stability of the enantiomers one quickly finds numerous methods. Picking a suitable procedure can be challenging, especially in fundamental research where new and undocumented systems are put forward. Fig. 2 gives an overview over the most commonly used techniques that are well accepted in the scientific community. Provided there are diastereotopic moieties present, a fast and reliable scouting method is dynamic NMR spectroscopy (DNMR) which can be used to estimate the enantiomerization time scale of the system at hand. If the rate of enantiomerization matches the timeframe of the NMR experiment, that is the inversion proceeds at an observable rate at a
coalescence temperature $T_c$ (see next section) within the temperature window of the instrument ($T_{\text{min}} < T_c < T_{\text{max}}$). DNMNR is the method of choice (1$\text{h} \leq t \leq 1$ hours – days). If the process is much faster than the NMR timescale ($T_{\text{max}} > T_c$), the conformers are short-lived and are not considered to be atropisomers ($t_{1/2} \approx$ seconds – minutes). If the inversion requires temperatures much higher than the instrument can provide ($T_c > T_{\text{max}}$), dynamic chromatography and dynamic chiroptical spectroscopy are likely candidates to obtain accurate enantiomerization rates and with that the thermodynamic parameters ($t_{1/2} \approx$ hours – days).

4. Dynamic NMR Spectroscopy (DNMR)

Dynamic NMR spectroscopy is one of the most prominent methods of determining the inversion barrier of atropisomers as it allows to do so directly from a racemic mixture. This makes this method particularly useful for enantiomerizations that occur rapidly at room temperature, in particular rotations around single bonds. Whereas in the other established methods the spectra show the time/temperature-dependent changes of the enantiomeric distribution in a racemization process, the spectra of the enantiomers are identical in DNMNR. It relies rather on enantiotopic nuclei or groups (usually protons) that are rendered diastereotopic due to the spatial configuration of the molecule. In a nutshell, the chiral configuration of the molecule locks two otherwise identical groups in place, which then experience a different chemical environment. The resulting anisochrony is then used as an internal gauge to observe the enantiomerization directly. A typical setup includes an NMR spectrometer, the chiral configuration of the molecule or group, and an NMR spectrometer with an internal gauge to observe the enantiomerization directly. The system possesses a bridge that shows diastereotopic protons due to the helical arrangement of the alkyl bridge. In this example, we will demonstrate the evaluation of a propyl bridged biphenyl (Fig. 3), which could not be separated into the isomers by DHPLC. The system possesses a bridge that shows diastereotopic protons due to the helical arrangement of the structure. The high-lighted protons can be used to determine the rate of inversion by measuring the peak broadening at half-height and subsequently determining the coalescence temperature. It is very important to calibrate the internal temperature of the sample beforehand, for which usually additives are used.

The determined coalescence temperature can now be used to calculate the barrier by using Eqn. (34):

$$k_{e,\bar{e}} = \frac{\pi \Delta v}{\sqrt{2}}$$  \(\Delta G_{e,\bar{e}} = RT_c \left[\ln \left(\frac{k_{e,T_c}}{h}\right) - \ln \left(k_{e,T_c}\right)\right]$$

$$\Delta G_{e,\bar{e}} = RT_c \left[\ln \left(\frac{k_{e,T_c}}{h}\right) - \ln \left(\frac{\Delta v}{\sqrt{2}}\right)\right]$$

If coupling occurs $k_{e,T_c}$ can be numerically approximated by Eqn. (35):

$$k_{e,T_c} = 2.22 \Delta v^2 + 6j^2$$

Simulating the spectra at variable temperatures (line shape analysis, LSA) gives access to the temperature dependent rate constants $k_{e,T_c}$, which in turn can be used to determine the thermodynamic parameters (Eyring plot, see Eqns (29)–(31) in section 2.6).

While being an elegant method, DNMNR requires the presence of diastereotopic protons, expensive hardware and usually prolonged measurement times to obtain precise spectra. The boiling points of the solvents often limit the window of observable enantiomerizations ($\Delta G^f$ typically 20–100 kJ mol$^{-1}$). Further requirements are temperature stability of the sample and a reasonably simple NMR spectrum without a multitude of coupling in the range of the diastereotopic protons.

4.1 Illustrated Example

Alkyl bridged biphenyls are well studied examples of atropisomers whose dynamic behavior has been studied in detail. Their relatively low lying barriers of inversion allow the kinetics to be studied by either DHPLC or DNMNR, depending among other things on the length of the alkyl bridge. In this example, we will demonstrate the evaluation of a propyl bridged biphenyl (Fig. 3), which could not be separated into the isomers by DHPLC. The system possesses a bridge that shows diastereotopic protons due to the helical arrangement of the structure. The high-lighted protons can be used to determine the rate of inversion by measuring the peak broadening at half-height and subsequently determining the coalescence temperature. It is very important to calibrate the internal temperature of the sample beforehand, for which usually additives are used.

The determined coalescence temperature can now be used to calculate the barrier by using Eqn. (34):

Fig. 3. Procedure to determine the theoretical coalescence temperature $T_c$. Determination of the line-width at half-height of the anisochronous signal (arrows and dotted lines, left). Plotting the obtained line-width versus the recorded internal temperature (right) and a subsequent Lorentzian fit allows the determination of the temperature at which the maximum line-width occurs. That temperature corresponds here to $T_c = 262.3 \pm 0.4$ K.
For the determination of the thermodynamic parameters, a range of rates at different temperatures is required. For the evaluation of the coalescence temperature, a set of spectra at various temperatures were already recorded, which can now be simulated by line-shape analysis (Fig. 4). The line-shape analysis gives access to the rates at these temperatures \(k_{\text{LSA}}\).

Once the rates are determined by LSA, an Eyring plot (see section 2.6) and a subsequent linear regression of the data can then be used to determine the thermodynamic parameters (Fig. 5). A word of caution: In almost all Eyring plots the experimental parameters \(\ln(\frac{k_{\text{LSA}}}{h})\) are used to determine the thermodynamic parameters, especially for entropy and enthalpy, should be treated carefully. This is the case for almost any Eyring regression.

Substitution into Eqns (29)–(31) gives:

\[
\Delta H^0 = -aR = -(-57912 \cdot 8.31 \cdot 10^{-3}) = 48.2 \text{ kJ mol}^{-1}
\]  

(37)

\[
\Delta S^0 = R \left[ b - \ln \left( \frac{aR}{T} \right) \right] = 8.31 \cdot \left[2.18 - 23.8\right] = -16.3 \text{ J mol}^{-1}
\]

(38)

\[
\Delta G^0_{298} = \Delta H^0 - T \Delta S^0 = 48.2 - (-0.0163 \cdot 298) = 53.1 \text{ kJ mol}^{-1}
\]

(39)

To compare with the value obtained directly from the coalescence temperature \(T_c = 262 \text{ K}\), the \(\Delta G_c^0\) value can be calculated at that temperature \(i.e. \Delta G_c^0_{262}\):

\[
\Delta G_c^0_{262} = \Delta H^0 - T \Delta S^0 = 48.2 - (-0.0163 \cdot 262) = 52.5 \text{ kJ mol}^{-1}
\]

(40)

which is in good agreement with \(\Delta G_c^0_{262}\) derived directly from \(k_{Tc}\) \(\Delta G_c^0_{Tc} = 0.2 \text{ kJ mol}^{-1}\).

5. Dynamic Chromatography

While DNMR allows a convenient access to the rate constant of enantiomerization \textit{in situ} directly from the racemate, it is best suited for fast enantiomerization processes (20–100 kJ mol\(^{-1}\)). Many of the more stable atropisomers are however beyond the instrumental limitations. For such systems one typically refers to either dynamic spectroscopy (which will be covered in the ensuing section) or dynamic chromatography\([6,16]\) which will be covered here.

Dynamic chromatography\([16c,17]\) has several advantages, the most prominent being that it allows the investigation of the rate of enantiomerization from a sample purified on-column, which excludes interference from chiral and achiral impurities. It is not necessary, contrary to the chiral spectroscopy methods, to obtain enantiomerically pure samples beforehand. Notably, the determination of the rate and with that the thermodynamic parameters require only the chromatogram. A typical setup includes besides a complete HPLC system a column oven with temperature control and rather expensive columns with a chiral stationary phase. Numerous methods that fall under dynamic chromatography exist, the most prominent being dynamic high pressure liquid chromatography (DHPLC), dynamic gas chromatography (DGC)\([5c]\) and dynamic supercritical fluid chromatography (DSFC)\([16d]\) and most of the mathematical background, advantages and shortcomings are at least comparable. Dynamic HPLC is by far the most common for systems such as ours and serves as the illustrative example. It typically allows the investigation of barriers in the range of 60–120 kJ mol\(^{-1}\) which translates into half lives of hours.

In principle chiral discrimination of the enantiomers is achieved by a chiral stationary phase on the column. Analogous to DNMR, peak coalescence can be observed by simultaneously resolving the enantiomers.
beforehand. Once resolution is achieved, the elution profiles are typically recorded over a range of temperatures, which is often limited by the boiling/freezing point of the solvent and the temperature stability of the column. Computer simulation of the elution profiles\textsuperscript{[35,10]} at these temperatures provides the individual rate constants \(k_i\) which in turn gives access to the thermodynamic parameters much as in the case of DNMR. The obtained data can be distorted, as in reality on-column enantiomerization does not proceed entirely in an achiral environment. The stationary phase is of course chiral which will have an influence on the determined barrier and the other thermodynamic parameters.\textsuperscript{[18]} The enantiomerization will also not, contrary to DNMR and dynamic chiroptical spectroscopy, be a uniform process which can be described with a single rate constant. The observed process is rather an average between rates of desorption and absorption, enantiomerization in solution (when desorbed) and in the solid state (when adsorbed). More modern methods that can overcome\textsuperscript{[10]} some of these limitations have demonstrated that these effects are however usually small. Further limitations include the choice of a suitable eluent with the appropriate polarity and solubilizing capability, the limited available temperature range, and the common requirement of developing a method of separation.

5.1 Illustrated Example

A dipiperidinyl-substituted butyl-bridged biphenyl from our group serves as a representative example.\textsuperscript{[16b,20]} After the development of a suitable HPLC method the resulting elution profile was investigated using the software DCXplorer by Trapp and coworkers. The key parameters that are determined (Fig. 7) are the separation of the peaks, the width and height of the plateau. From that the corresponding rate constant \(k_e\) was determined computationally.

The example shown in Fig. 7 yields \(k_{298} = 1.24 \cdot 10^{-3} \text{ s}^{-1}\). Substitution into Eqn. (23) gives Eqn. (41):

\[
\Delta G^\ddagger_{298} = RT \left[ \ln \left( \frac{k_e}{k_2} \right) - \ln \left( k_e \right) \right]
\]

\[
= 8.31 \cdot 10^{-3} \cdot 298 \left[ \ln \left( 1.38 \cdot 10^{-23,298} \right) - \ln \left( 1.24 \cdot 10^{-3} \right) \right] = 89.6 \text{ kJ mol}^{-1}
\]

6. Dynamic Chiroptical Spectroscopy

The third prominent and probably also the oldest method to determine inversion barriers is to observe the degeneration of a chiroptical signal.\textsuperscript{[21]} An enantiomerically enriched sample that is optically active...
decays exponentially into the chiroptically mute racemate[22] which can be followed by dynamic chiroptical spectroscopy. Polarimetry and more recently dynamic circular dichroism (DCD) are the analytical methods used. The observed barriers are typically in a range of 80–180 kJ mol\(^{-1}\).[8] Below 80 kJ mol\(^{-1}\) the optically active sample degenerates before the measurement to a considerable amount. Above 180 kJ mol\(^{-1}\) the time required to record the process is significant (> one week). Such samples are generally considered to be stable enantiomers at room temperature. A typical setup includes hardware for the separation of the enantiomers (often a HPLC setup including a column with a chiral stationary phase) and a spectrometer (CD or optical activity) with a heating block to change the temperature of the sample.

The process observed is the conversion of an enantiomer into the racemic mixture (racemization). Careful, the obtained value for \(k\) is \(k_{rac}\) not \(k\) which is used for the determination of thermodynamic parameters. But as \(2k = k_{rac}\), the conversion is straightforward. Rearranging Eqn. (16) gives:

\[
\ln \frac{\Delta A}{\Delta A_0} = -2k t \quad \text{ln}(\Delta A)
\]  

(45)

as the relationship between the CD signal and the initial concentration is linear. \(\Delta A\) and \(\Delta A_0\) are the difference in absorption at \(t = 0\) or \(t\), respectively.

Plotting \(t\) against \(\ln(\Delta A)\) gives access to \(k_{rac}\) directly from the slope (\(a = -2k_{rac}\)). The obtained \(k_{rac}\) can then be used to calculate the free Gibbs energy of racemization \(\Delta G^{\circ}_{sect}\) by using the rearranged Eyring equation:

\[
\Delta G^{\circ}_{sect} = \Delta G^{\circ}_{sect} - RT \ln \left( \frac{k_{rac}}{k_B T} \right) = \Delta G^{\circ}_{sect} - \frac{RT}{k_B T} \ln \left( \frac{k_{rac}}{k_B T} \right)
\]  

(46)

Fig. 8. Eyring plot of \(\ln(k_{rac}/T)\) vs. 1/T of all evaluated elution profiles and subsequent linear regression \(y = ax + b\) with slope \(a = -5051 \pm 112\) and intercept \(b = 4.58 \pm 0.4\). Shading indicates the 95% confidence interval, error bars the standard deviation at each temperature.

6.1 Illustrated Example

Along the numerous chiral systems we developed, the ‘Geländer’-oligomer[24] is an ideal candidate to illustrate the use of dynamic chiroptical spectroscopy. The hexaphenyl structure exists exclusively as two helical enantiomers and was suited to study enantiomerization dynamics involving the rotation of more than one single bond in the backbone. The barrier of helix inversion was found to be too elevated to be measured by DNMR leaving two options: DHPLC or DCD. After an extensive screening of conditions we achieved the separation as shown in Fig. 9, top. While both enantiomers are resolved as peaks, the missing baseline separation prevented the determination of the rate constant by DHPLC. However, the resolution was good enough to obtain the enantiomers in >95% ee directly off the column which allowed the decay of the chiroptical information to be studied over time (Fig. 9, middle). The logarithm is particularly sensitive to small numbers close to zero which results in scattered data at the end of the experiment. It is therefore advisable to use strong CD bands that deviate significantly from zero to obtain a large linear segment to extend the dataset for the linear fit. In the example described here most of the data points were suitable leading to

\[
a = -1.29 \cdot 10^{-4} = -2k_e
\]  

(47)

\[
k_e = \frac{1}{a} = 8.09 \cdot 10^{-4} \text{ s}^{-1} = \frac{1}{2} \cdot \frac{4.67 \cdot 10^{-5}}{298 \cdot 8.31 \cdot 10^{-3}} = 2 \cdot \frac{6.31 \cdot 10^{-34}}{6.63 \cdot 10^{-34}} = 2
\]  

(48)

Substitution into Eqn. (23) gives Eqn. (49):

\[
\Delta G^{\circ}_{298} = RT \ln \left( \frac{k_{ee}}{k_{ee}} \right) = \Delta G^{\circ}_{298} - \frac{RT}{k_B T} \ln \left( \frac{k_{ee}}{k_{ee}} \right) = \frac{1.39 \cdot 10^{-22}}{6.63 \cdot 10^{-34}} - \ln \left( \frac{6.47 \cdot 10^{-5}}{6.47 \cdot 10^{-5}} \right) = 96.9 \text{ kJ mol}^{-1}
\]  

(49)

Like in the example for DHPLC, measuring decay profiles for multiple temperatures would allow the remaining thermodynamic parameters (entropy and enthalpy) to be obtained.

7. Errors and Quality of Fits

The presentation of data in publications is accompanied by its estimation of the error. It describes the robustness of the data and their evaluation and indicates precision and uncertainties. However, a proper error analysis is time consuming and in many cases very challenging. This section explains some of the often encountered variables used in error determination.
and presents methodologies for a solid and (relatively fast) error estimation. We will refrain from an in-depth discussion and refer the reader to the literature.

### 7.1 Evaluating the Quality of Linear Fits

Linear regression is essentially a simpler, special case of the nonlinear counterpart and nonlinear regression programs can be used to fit linear models. But fitting linearized data can be delicate. One of the most fundamental drawbacks of linearization procedures is that linear regression calculations are only valid when the experimental uncertainty of y values is not related to the values of x or y. This is not always the case because linearization procedures distort experimental errors and may lead to their over- or underrepresentation in frontier values of x and y. This effect is even more severe in transformations combining x and y values. While the distortion can be usually neglected in first order processes such as enantiomerizations – the errors in time or temperatures are usually relatively low – it is nevertheless important to be aware of the issue, especially when dealing with aggregation or similar phenomena that follow complex non-linear behavior.

After the fitting procedure, the goodness of fit has to be assessed. Often, problems with the fit can be detected visually by superimposing the graph of the fit with the data points. Important tests include the use of residual plots, the Wald-Wolfowitz test, the reduced coefficient of determination \( R^2 \) and reduced Chi-Square \( \chi^2 \).

A residual plot shows the difference between the measured and the calculated data as a function of x. Ideally, the residuals account for experimental error only, that is, the data is appropriately represented by the formula. Thus, a trend in this plot would suggest a bad fit; the residual plot should show a random distribution of negative and positive residuals. Clustering of the positive and negative residuals is indicative of systematic differences between the data and the predicted values from the curve.

The runs test (Wald-Wolfowitz Test) is a quick and solid means to detect whether the regression curve deviates systematically from the data. A run can be defined as a consecutive series of points where the residuals have the same sign. If \( N \) is the total amount of data points with \( N_a \) points above the curve and \( N_b \) points below, the expected number of runs equals \( \left( \frac{2N_aN_b}{N_a+N_b} \right) + 1 \). Critical values tables for numbers of runs depending on \( N_a \) and \( N_b \) are readily available to affirm or reject randomness of the residuals.

The \( R^2 \) value is the coefficient of determination and often automatically determined by the statistical evaluation software. It is used to account for the degree to which the variance of the ‘dependent’ variable is explained by the ‘independent’ variable:

\[
R^2 = 1 - \frac{\sum(y - y_{\text{fit}})^2}{\sum(y - y_{\text{mean}})^2}
\]

\( R^2 \) values range between 0 and 1. For example, an \( R^2 \) value of 0.997 means that 99.7% of the variation of the dependent variable is caused or explained by the variation of the independent one. The more parameters are contained in the formula, the closer the value of \( R^2 \) to 1, independently from the contribution of the parameters to
the model. The adjusted $R^2$ value, $(\tilde{R}^2)$, corrects this bias by adjusting the number of parameters in relation to the number of data points. It is usually smaller than the regular $R^2$ value:

$$\tilde{R}^2 = 1 - \frac{(n-1) \sum (y-y_{\text{fit}})^2}{(n-K) \sum (y-y_{\text{mean}})^2}$$  \hspace{1cm} (51)

with $n$ being the number of data points and $K$ the number of parameters fit by the regression; the expression $(n-K)$ gives the number of degrees of freedom of the regression.[26,27]

Reduced Chi-square $\chi^2_{\text{red}}$ is best described by first describing Chi-square $\chi^2$. In linear regression $\chi^2$ is a measure of how well the observed data is consistent with the theoretical (calculated) data. $\chi^2$ is expected to be of order $n$ (number of data points) if the fit is good. A value that is much greater than $n$ is indicative of a poor fit. However, $\chi^2$ can only be used if the theoretical values and the standard deviation are known, which is not often the case. The difference between the reduced chi-square $\chi^2_{\text{red}}$ and chi-square $\chi^2$ is that while the latter is referenced to the standard deviation, the former is referenced to the degrees of freedom $(n-K)$:

$$\chi^2_{\text{red}} = \frac{\chi^2}{n-K}$$  \hspace{1cm} (52)

In theory, $\chi^2_{\text{red}}$ has the expected value of 1, where a value larger than 1 indicates a wrong model and a value smaller than 1 overparametrization of the model.[28] Careful, sometimes programs use a different definition for $\chi^2_{\text{red}}$ which can lead to confusion.[29]

### 7.2 Estimation of Uncertainties (Error Calculation)

The accuracy and reliability of the results comprises two components of uncertainty, namely the experimental repeatability uncertainty and the fitting process (data analysis) uncertainty.[30]

Assuming no systematical errors in the analysis,[31] an estimation of uncertainty regarding the repeatability of the experimental data can be obtained by repeating the experiment and analyzing the data in (near-)identical manner six to eight times.[23] Because this is not always feasible for economic reasons or time constraints a minimum of three experimental cycles is recommended. This gives a good idea of the reliability of the results and enables a conservative estimation of the uncertainties. For each set of three data points (for instance at a certain temperature) the corresponding standard-deviation can be calculated, if $x$ can be determined accurately. The standard deviation can then be used as error bar for that set of points, which in turn allows the computation of confidence bands and so on.[30] Especially when $x$ denotes temperature or time the standard deviation provides an appropriate uncertainty of $y$.

Although data may be linearized (as is the case for CD decay curves which are then plotted logarithmically), the process under investigation is still nonlinear and no straightforward procedure exists to estimate the uncertainty during the fitting process. Most programs display the 95% confidence interval or error estimates on the fitted data.[25,33] These asymptotic standard errors assume, among other things, that the physical property $y$ is the only source of experimental error and do not account for asymmetrical errors. However, asymptotic standard errors are still precise enough to qualitatively evaluate the results of the fit.[20] Very large asymptotic standard errors may be indicative of a variety of problems with the process investigated, ranging from lack of precision of the measurements to wrong choice of model and might even indicate the process not taking place. When errors need to be known more precisely, the described disadvantages can be avoided by using the Monte Carlo simulation method.[34] The underlying concept of Monte Carlo simulations is to simulate data sets with different random scatter for each set. After fitting each simulation, the distribution of the fitted parameters is used to create confidence intervals. The implementation of Monte Carlo simulation generally consists of five steps which are demonstrated here by estimating the uncertainty of the thermodynamic parameters obtained by an Eyring plot:[26b,34]

1) Compute an optimized data set. In our example, values for $\Delta H^\text{‡}$ and $\Delta S^\text{‡}$ were determined (Fig. 10, top) which can be used to recursively calculate the corresponding rates $k_e$ at $n$ (here: $n = 6$) given temperatures $T$ by Eqn. (27) (Fig. 10, middle). Since these rates are calculated, the resulting linear regression of these $n = 6$ values yields a perfectly straight line.

2) Add (Fig. 10, bottom) scatter $\sigma$ (typically derived from a Gaussian distribution) to each of the $n = 6$ data points such that $T' = T + \sigma$, and $k_i' = k_i + \sigma$, using the Box-Muller- or the Marsaglia polar method, for example. In other words, to each temperature $T$ (x-axis) and rate $k_i$ (y, although the y-axis in Eyring-plots displays $\ln(y)$) is added a certain, randomly chosen value $\sigma$ (positive or negative) that causes the data points to scatter around the optimized data set in two dimensions; this should account for most, if not all, instrumental uncertainties in the original $n = 6$ measurements.

3) Fit the simulated data and collect the optimized values of the adjustable parameters from which an Eyring plot delivers, $\Delta H^\text{‡}_\text{Eyring}(\sigma_1)$, $\Delta S^\text{‡}_\text{Eyring}(\sigma_1)$ and $\Delta G^\text{‡}_\text{Eyring}(\sigma_1)$.

4) Steps 2) and 3) are repeated $m \geq 1000$ times (here 2500 times). For each of these $m$ simulations for $n$ data points an Eyring plot delivers $\Delta H^\text{‡}_\text{Eyring}(\sigma_m)$, $\Delta S^\text{‡}_\text{Eyring}(\sigma_m)$ and $\Delta G^\text{‡}_\text{Eyring}(\sigma_m)$. By definition, the set of the calculated values corresponds to the distribution function of the corresponding parameter. These distributions can be visualized by plotting a histogram with the $\Delta H^\text{‡}_\text{Eyring}(\sigma_m)$, $\Delta S^\text{‡}_\text{Eyring}(\sigma_m)$ and $\Delta G^\text{‡}_\text{Eyring}(\sigma_m)$ values in the x-axis and their frequency (i.e. the number of times this value was found in the $m$ repetitions) in the y-axis.
5) Determine the appropriate percentile values (for example 97.5 and 2.5 for the 95% confidence interval) for every parameter; these values constitute the frontiers of the confidence interval (i.e. the confidence interval is contained between these two frontier values). Returning to the histogram, this is identical to removing the \( \Sigma y=25 \) lowest and \( \Sigma y=25 \) highest results. The remaining values are the 95% confidence interval. Keep in mind that these intervals need not necessarily be symmetrical.

A remark concerning confidence intervals: a confidence interval of 95% does not mean that the true value lies in the confidence interval with 95% probability; this is a measure of final precision, but the actual value of the parameter is unknown (thus the estimation procedures). Also, once the confidence intervals are established, it is not a matter of probability whether the interval contains the true parameter value. It either does or it does not. Confidence intervals are rather a measure of the reliability of the estimation procedure itself. A useful interpretation of confidence intervals is the following: the event of the true value lying outside the 95% confidence intervals has the probability of happening 5% (or less) by chance. [26, 35]

8. Comparison of Methods

Within this manuscript we have presented key aspects and methods to determine the thermodynamic parameters of enantiomerization processes, especially the racemization barrier. Among the most frequently used techniques are DNMR, DHPLC and dynamic chiroptical spectroscopy. Each of these methods are especially suited for a specific range of enantiomerization barriers, and are summarized (Table 1) in a simplified table in analogy to Wolf.[36] Depending on the nature of the system under investigation, any of these methods will provide a relatively fast access to the thermodynamic parameters including the barrier of inversion.

It is our sincere hope that this tutorial will be a useful entry point into the fascinating world of enantiomerizations for the often reluctant organic chemist and will provide a sound basis both mathematically and experimentally to bring many more of these dynamic systems to light.

---

**Table 1. Comparison of covered methods to determine the thermodynamic parameters of enantiomerization based on Wolf.[36]**

<table>
<thead>
<tr>
<th>Method</th>
<th>DNMR</th>
<th>DHPLC</th>
<th>Dynamic Chiroptical Spectroscopy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principle to obtain rate constants ( k_{eq} )</td>
<td>Coalescence of NMR resonances is simulated by line-shape analysis.</td>
<td>Plateau-formation due to on column enantiomerization. Subsequent simulation of the profiles.</td>
<td>Observation of the decay of a chiroptical signal at various temperatures.</td>
</tr>
<tr>
<td>Barrier–range / ( \text{kJ mol}^{-1} ) (half life range)</td>
<td>20–130 (min – hours)</td>
<td>60–120 (hours)</td>
<td>80–180 (hours – days)</td>
</tr>
<tr>
<td>Requirements</td>
<td>Sufficient resolution of the anisochronous signals</td>
<td>Suitable method of separation</td>
<td>Semipreparative amounts of pure or enriched enantiomer.</td>
</tr>
<tr>
<td>Sample amount</td>
<td>1 – 10 mg</td>
<td>&lt;10 µg</td>
<td>0.5 – 10 mg</td>
</tr>
<tr>
<td>Typical uncertainty</td>
<td>± 0.2 (LSA)</td>
<td>± 0.5</td>
<td>± 0.5</td>
</tr>
<tr>
<td>Chiral impurities</td>
<td>Only if in observed shift region</td>
<td>Are usually removed during elution. Chiral stationary phase might interfere.</td>
<td>Interferes.</td>
</tr>
</tbody>
</table>

**Definitions**

\( X_t \) Variable \( X \) at a specific temperature \( T \), e.g. \( k_{eq} \)

\( X_{eq} \) Variable \( X \) as a function of temperature \( T \) (temperature range), e.g. \( k_{eq} \)

\( A \) and \( B \) Enantiomers of each other. Careful: \( A \) describes also the prefactor in the Arrhenius equation

\( A^\dagger \) Transition state between \( A \rightarrow B \)

\( k_{ac} \) Rate of racemization, i.e. \( 2A \rightarrow A + B \), in s\(^{-1}\)

\( k_e \) Rate of enantiomerization, i.e. \( A \leftrightarrow B \), in s\(^{-1}\)

\( k_{acB} \) Rate of the irreversible transformation \( A \rightarrow B \), in s\(^{-1}\)

\( k_{acA} \) Rate of the irreversible transformation \( B \rightarrow A \), in s\(^{-1}\)

\( [A]_t \) Concentration of \( A \) at time \( t \)

\( [A]_0 \) Concentration of \( A \) at time \( t=0 \)

\( \Delta G_{rac}^\circ \) Standard Gibbs free energy for racemization, in kJ mol\(^{-1}\)

\( \Delta G_{e}^\dagger \) Gibbs free energy of activation for enantiomerization, i.e. \( A \rightarrow B \), often called the barrier of racemization/inversion. Often also written as \( \Delta G_{rac}^\circ \) in kJ mol\(^{-1}\)

\( \Delta H_{e}^\dagger \) Enthalpy of activation for enantiomerization, i.e. \( A \rightarrow B \), in kJ mol\(^{-1}\)

\( \Delta H_{rac}^\circ \) Enthalpy of racemization, in kJ mol\(^{-1}\)

\( \Delta S_{e}^\dagger \) Entropy of activation for enantiomerization, i.e. \( A \rightarrow B \), in J mol\(^{-1}\) K\(^{-1}\)

\( \Delta S_{rac}^\circ \) Entropy of racemization, in J mol\(^{-1}\)

\( \Delta \) Difference in chemical shift of two diastereotopic moieties without exchange in Hz

\( J_{AA'} \) Coupling constant between diastereotopic nuclei \( A \) and \( A' \) in Hz

\( k_B \) Boltzmann constant, \( 1.380662 \times 10^{-23} \) J K\(^{-1}\)

\( h \) Planck’s constant, \( h = 6.626176 \times 10^{-34} \) J s

\( R \) Universal gas constant, \( R = 8.31446 \times 10^{-3} \) kJ K\(^{-1}\) mol\(^{-1}\)

\( T \) Temperature in Kelvin

\( T_c \) Coalescence temperature in Kelvin.

\( E \) Activation energy in kJ mol\(^{-1}\)

\( W \) Number of possible microstates in a system

**Acknowledgements**

The authors acknowledge financial support by the Swiss National Science Foundation (SNF).

Received: December 7, 2015

---


