

8	Engineering enzyme selectivity	203
8.1	Introduction	204
8.1.1	Stereoisomers	204
8.1.2	Stereoisomeric purity	205
8.2	Product selectivity	207
8.3	Substrate selectivity	209
8.3.1	Measuring selectivity in kinetic resolutions	211
8.4	Molecular basis of selectivity	213
8.5	Engineering selectivity (incomplete)	215
8.6	Specialist and generalist enzymes	215
8.7	Engineering the substrate range of an enzyme	217
	References	220
	Problems	221
	Supporting Information	221

8 Engineering enzyme selectivity

© 2023 Romas Kazlauskas

Summary, text to be added later

Learning goals

- Selectivity is the ability of proteins to distinguish between different molecules. The selectivity value, S , is the ratio of the preferred to non-preferred molecule. This value depends on comparison molecules. Selective reactivity may involve discrimination between two substrates or between two products.
- Stereoselectivity is a type of selectivity where the molecules to be distinguished are stereoisomers; enantioselectivity is a type of stereoselectivity where the stereoisomer to be distinguished are enantiomers. Many pharmaceuticals require pure enantiomers and enzyme-catalyzed reactions are a good way to make these pure enantiomers.
- Discrimination between two substrates occurs when one substrate reacts faster than the other. This substrate selectivity depends on ratio of k_{cat}/K_M for the two substrates. Tighter binding allows a substrate to occupy a larger fraction of active sites, while faster reaction allow the substrate to react faster.
- When there are equal amounts of the two substrates, the initial ratio of substrates that react or products that form corresponds to the selectivity. At higher conversions, the relative amounts of substrate that reacts or product that forms differs from the selectivity because the amounts of the two substrates is no longer equal.
- Discrimination between two products occurs when a single substrate forms two products. This product selectivity depends only on the relative values of k_{cat} for the two products. The relative amount of the two products is constant throughout the reaction and corresponds to the selectivity.

8.1 Introduction

Selectivity, S , is the ability to distinguish between alternatives. This ability is the key to the value of many proteins. For example, the cardiovascular antibody drug abciximab selectively binds to a receptor involved in platelet aggregation to inhibit formation of blood clots.^[1] In another example, a transaminase with high selectivity for the (S)-amine yields the required stereoisomer of a cardiovascular drug precursor with high purity.^[2]

Binding selectivity is the ability to bind one ligand more tightly than other one and was covered in Chapter 6. This chapter focusses on reaction selectivity.

Reaction selectivity is the ability to react with one substrate more quickly than another (substrate selectivity) or form one molecule more quickly than another (product selectivity). The classification as substrate or product selectivity depends on whether there are multiple possible substrates or multiple possible products.

Reaction selectivity stems from differences in reaction rates, which originate in differences in transition state energies for the competing reactions. The ratio of rates is the selectivity for that pair of molecules, eq. 8.1.

$$S_{reaction} = \frac{rate_{fast}}{rate_{slow}} \quad (8.1)$$

The words non-selective (or promiscuous), selective and highly selective describe degrees of selectivity. Avoid using the word specific to mean highly selective because specific has a different meaning in chemistry. The reaction mechanism of a selective reaction favors the formation of a particular isomer, but both isomers can form. For example, epoxidation of norbornene (bicyclo[2.2.1]hept-2-ene) favors formation of the *exo* diastereomer (oxygen close the one-carbon bridge) over the *endo* diastereomer (oxygen close to the two-carbon bridge), Fig. 8.1a. The selectivity originates the different steric hindrance between the two sides and varies with the oxidant used. A highly selective reaction may form only one detectable product, but the the reaction mechanism could also form the other one. In contrast, specific reactions are those that require formation of a particular isomer for mechanistic reasons. For example, epoxidation of *cis*-2-butene can only form the *cis*-epoxide, Fig. 8.1b. Likewise, epoxidation of *trans*-2-butene can only form the *trans*-epoxide. The reaction is a concerted addition of oxygen to the double bond without an intermediate that permits rotation of the carbon-carbon bond. The reactions in this chapter all refer to selectivity.

8.1.1 Stereoisomers

Isomers are non-superposable molecules with the same molecular formula. Isomers can differ in their atom connectivity (constitutional isomers) or only in shape (stereoisomers), Fig. 8.2. Classifying a person as a father, son, brother or uncle depends on who the person is being compared to. In the same way, isomerism classifies relationships between molecules. A given molecule may be a diastereomer, an enantiomer or a constitutional isomer depending on the comparison molecule. For example, one may recognize

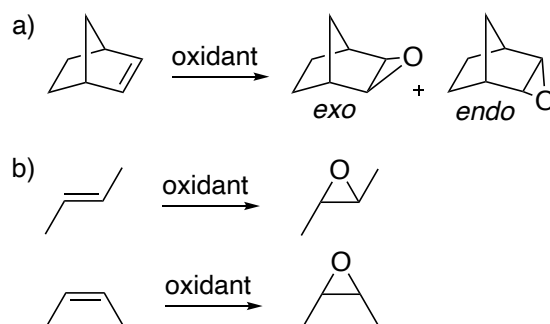


Figure 8.1. Selective reactions favor one possible product, but specific reactions form only one product because the reaction mechanism make formation of the other one impossible. a) Epoxidation of norbornene is a selective reaction. The oxygen atom can add to either face of the double bond. b) Epoxidation of 2-butene is a specific reaction: *cis*-2-butene yields only the *cis*-epoxide, while *trans*-2-butene yields only the *trans*-epoxide.

that a molecule is chiral, but until one has a comparison molecule, one cannot classify the relationship between the molecules.

The concept of a stereogenic unit is also useful. A stereogenic unit is the part of the molecule that makes stereoisomers possible. Chiral molecules must contain at least one stereogenic unit. For example, the stereogenic unit in *L*-alanine is the α -carbon atom containing four different substituents: HOOC- , $\text{H}_2\text{N-}$, $\text{H}_3\text{C-}$, and H . Occasionally one may encounter the term regioisomer, which refers to a special type of constitutional isomers. Regioisomers can form when a chemical reaction has different orientations or sites to choose from. For example, electrophilic aromatic substitution of bromobenzene can yield *ortho*-, *meta*-, or *para*-substituted products. These isomeric products are regioisomers.

Researchers often use the high stereoselectivity of enzymes, especially their high enantioselectivity to prepare pure stereoisomers. Most of the examples in this chapter involve enzymes distinguishing between stereoisomers.

8.1.2 Stereoisomeric purity

There are two commonly used measures for stereoisomeric purity: the ratio of stereoisomers and the stereoisomeric excess. The enantiomer ratio, *er*, or diastereomer ratio, *dr*, is simply the relative amounts of the two stereoisomers.

$$er \text{ or } dr = \frac{\text{major stereoisomer}}{\text{minor stereoisomer}}$$

For example, a 9:1 mixture of *cis*- and *trans*-2-butene has a diastereomeric ratio of 9. Analytic methods like HPLC or GC with columns containing a chiral stationary phase reveal the amounts of each enantiomer.

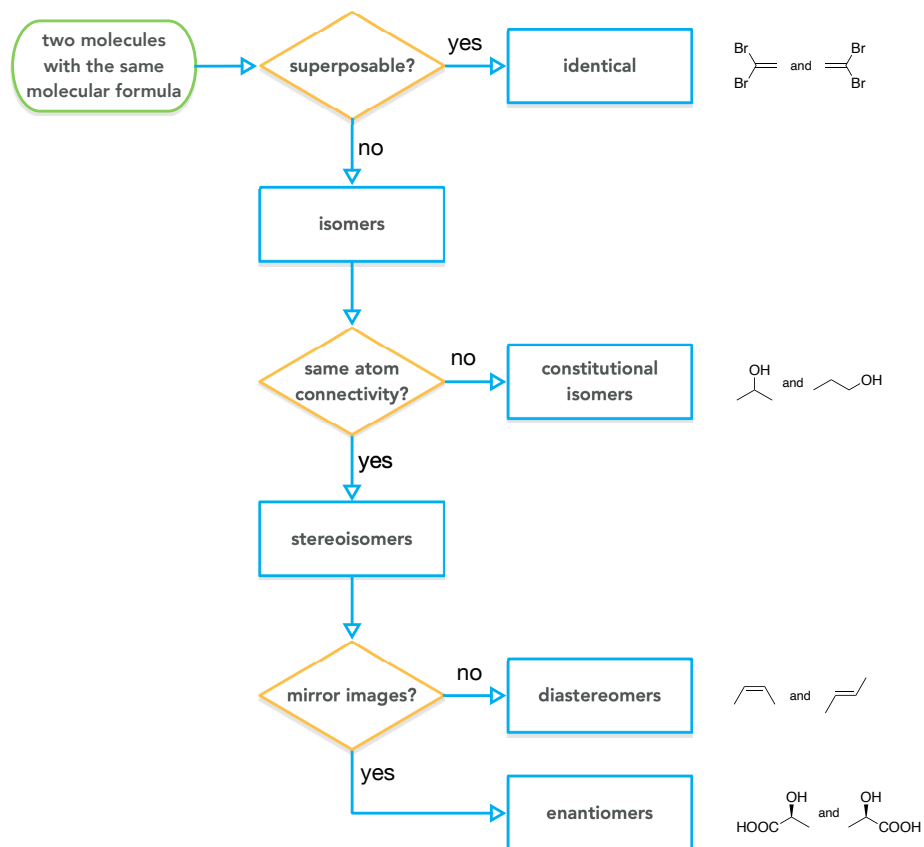


Figure 8.2. Flowchart to identify common relationships between pairs of molecules. Constitutional isomers differ in their atom connectivity, while stereoisomers differ only in their shape. Stereoisomers can be either diastereomers or enantiomers.

Another measure of purity is enantiomeric excess, *ee*, usually expressed as a percent instead of a number between 0 and 1. Enantiomeric excess, is the excess of the major enantiomer over the minor enantiomer in the sample, eq 8.2. For example, a 9:1 mixture of two enantiomers has an 80% *ee*. Diastereomeric excess, *de*, is defined similarly.

$$ee \text{ or } de = \frac{major - minor}{major + minor} \quad (8.2)$$

This unusual way to express purity stems from the historic use of optical rotation to measure enantiomeric purity. Enantiomers rotate plane-polarized light in opposite directions. A racemic sample shows no rotation because the opposite rotations of the two enantiomers cancel out. A racemic sample has an *ee* of 0%. If one enantiomer is present in excess, then the excess amount of that enantiomer causes a rotation. The enantiomeric excess in percent is the size of that rotation as a fraction of the rotation for the pure enantiomer.

To convert enantiomeric excess to amounts of the individual enantiomers, one can imagine the sample to consist of an excess amount of the major enantiomer, *ee*, plus an amount of racemate $1 - ee$. The total amount of major enantiomer is the excess amount plus half of the racemate, eq. 8.3.

$$major_{total} = major_{excess} + \frac{racemate}{2} = ee + \frac{1 - ee}{2} \quad (8.3)$$

The total amount of the minor enantiomer is half of the amount of the racemate, eq. 8.4.

$$minor_{total} = \frac{racemate}{2} = \frac{1 - ee}{2} \quad (8.4)$$

For example, a sample with an enantiomeric purity of 80% *ee* consists of 80% major enantiomer in excess plus 20% racemate. Thus, the sample contains 90% of the major enantiomer and 10% of the minor enantiomer.

8.2 Product selectivity

Product selectivity occurs when an enzyme converts a substrate into several possible products and favors one or more of the possibilities. For example, an enantioselective transaminase catalyzes the key step in the synthesis of sitagliptin, a type 2 diabetes drug^[3], Fig. 8.3. The reaction is highly selective for one of the two possible enantiomers. Enantioselectivity is a particularly useful type of selectivity since pure enantiomers are difficult to make with traditional chemical methods. Enantiomers can differ in their biological effects, so most chiral drugs are marketed as single enantiomers.

Product selectivity depends on the ratio of k_{cat} values for the two products, eq. 8.5. The subscript f refers to the faster-reacting substrate and subscript s refers to the slower reacting substrate. There is no contribution from binding, K_M , because there is only one

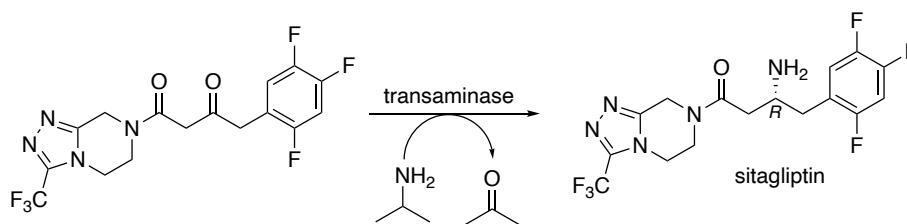


Figure 8.3. An enantioselective transaminase converts the achiral starting ketone into the (*R*)-amine. This reaction is an example of product enantioselectivity since the other possible product, the enantiomeric (*S*)-amine (not shown), forms only in trace amounts.

substrate, so no competition is possible, Figure 8.4. The Gibbs energy difference between the two transition states is proportional to the natural logarithm of the selectivity, eq. 8.6.

$$S_{product} = \frac{(k_{cat})_f}{(k_{cat})_s} \quad (8.5)$$

$$\Delta\Delta G_{f-s} = -RT\ln(S_{product}) \quad (8.6)$$

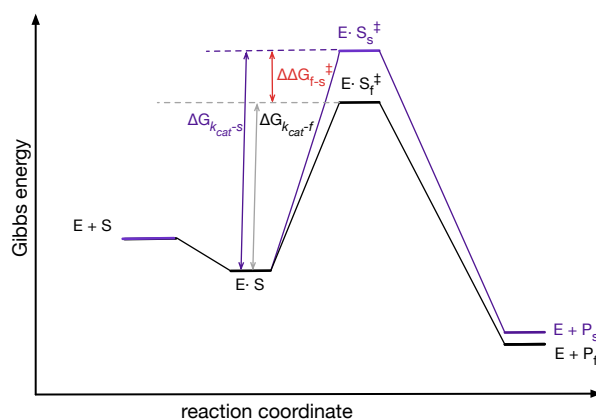


Figure 8.4. Gibbs energy diagram for product selectivity. The enzyme can convert a single substrate into one of two products. One product, P_f , forms faster because the transition state to form that product is lower in energy than the transition state to form the other product, P_s . The difference in Gibbs energy for the two transition states is the origin of the selectivity.

To measure product selectivity one measures the relative amounts of the products formed. The ratio of the products corresponds to the selectivity. This ratio remains constant as the reaction proceeds because the starting material never changes; it just decreases in amount.

8.3 Substrate selectivity

Substrate selectivity occurs when two substrates compete in the same solution. An enzyme encounters similar substrates, but reacts with some of them more rapidly than others. Proteases are a good example of substrate selectivity. A peptide contains multiple peptide links each of which is a different potential substrate. Proteases often favor hydrolysis of some of these links over others. For example, trypsin favors hydrolysis of peptide links after lysine or arginine residues, 8.5a. The example peptide contains fifteen amide links (fifteen possible substrates), but trypsin cleaves only two of these due to its substrate selectivity.

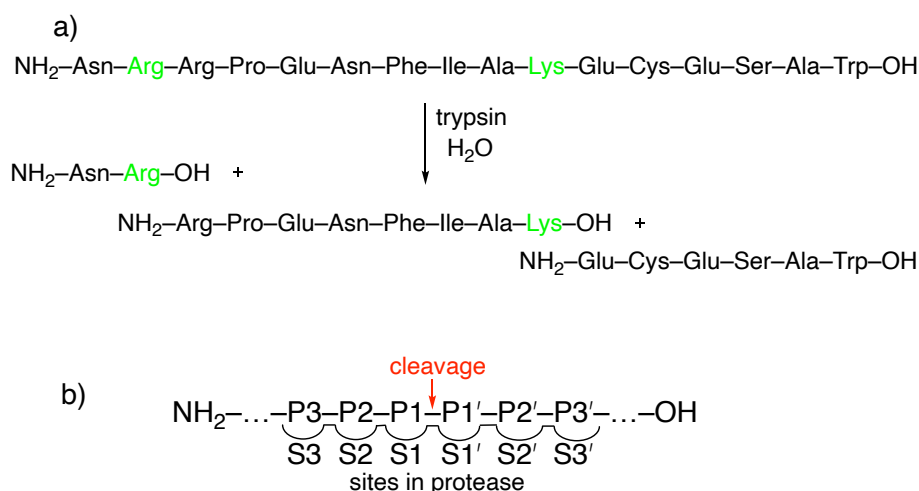


Figure 8.5. Proteases are often selective in which peptide links they cleave. a) Trypsin cleaves only two of the fifteen amide links in peptide shown. Trypsin is selective for amide links on the C-terminal side of lysine and arginine residue, except when followed by proline. b) The binding sites in the protease are numbered to indicate their distance from the catalytic site. The sites S1, S2, S3, etc. bind the N-terminal part of the peptide, while the sites S1', S2', S3', etc. bind the C-terminal part of the peptide. The peptide residue naming matches the names of the binding sites: P1, P2, P3, etc. in the direction of amino terminus of the peptide and P1', P2', P3', etc. in the direction of the carboxy terminus. The peptide residue naming changes as different amide bonds are positioned at the cleavage site.

Protease nomenclature describes how the peptide orients in the protease when cleavage occurs,^[4] Fig. 8.5b. The numbering of the binding sites, S, indicates their distance from the catalytic site where cleavage occurs: ...S3-S2-S1-S1'-S2'-S3'... The sites numbered without primes bind the N-terminal part of the peptide, while the sites numbered with primes bind the C-terminal part of the peptide. The sites bind both the side chains and main chain of the peptide. The catalytic site lies between S1 and S1' and cleaves the peptide bond connecting the two residues bound in those sites. The numbering of the enzyme sites is fixed by the location of the catalytic residues. The numbering of the

peptide substrate varies when different amide links are positioned at the cleavage site. The peptide numbering matches the binding sites at which it binds: $\text{NH}_2 \dots \text{P3-P2-P1-P1}'\text{-P2}'\text{-P3}' \dots \text{COOH}$. Trypsin is selective for peptide links with lysine or arginine at the P1 position, except when the P1' amino acid is proline. The proline at P1' presumably disrupts catalytically productive binding.

Substrate selectivity originates from both the ability of the enzyme to bind the good substrate and also to catalyze a reaction on the good substrate. The good substrate must succeed at both, while the poor substrate may fail at either step. Poor substrates either do not bind to enzyme or bind in a way that does not lead to reaction. The selectivity between the two depends on both k_{cat} and K_M , eq. 8.7, where subscript f refers to the faster-reacting substrate and subscript s refers to the slower reacting substrate.

$$S_{\text{substrate}} = \frac{(k_{\text{cat}}/K_M)_f}{(k_{\text{cat}}/K_M)_s} \quad (8.7)$$

The tighter binding substrate will occupy a larger fraction of the enzyme active sites, while the faster catalytic step will convert more of that substrate to product. The Gibbs energy diagram, Figure 8.6, shows these two contributions to substrate selectivity graphically. In this diagram, both k_{cat} and K_M favor the faster-reacting substrate, but it is also possible that they favor different substrates, in which case the competing effects lead to a lower net selectivity.

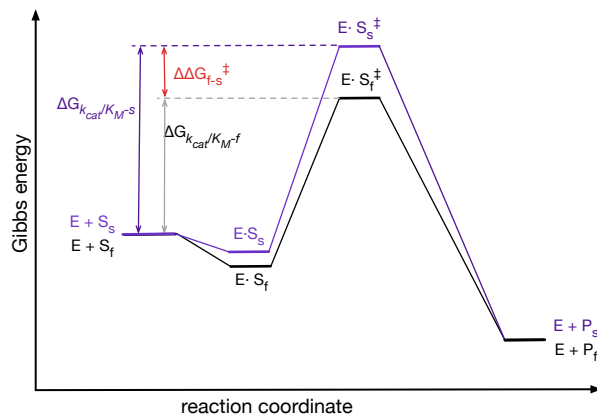


Figure 8.6. Gibbs energy diagram for two competing substrates in an enzyme-catalyzed reaction. Both tighter binding and faster reaction contribute to the selectivity. S_f = faster-reacting substrate; S_s = slower-reacting substrate.

For the trypsin example, the different K_M 's refer to the binding of different amide bonds in the catalytic site. To estimate K_M for the different amide bonds one could measure K_M using substrate analogs that contain only one type of amide bond.

A common special case of substrate selectivity occurs when the two competing substrates are enantiomers. In this case the reaction is a kinetic resolution. The word reso-

lution refers to a separation of enantiomers, while kinetic specifies that the separation relies on differences in the rates of reaction. For example, the manufacture of diltiazem, a calcium-channel blocker uses a kinetic resolution to make an enantiomerically-pure precursor,^[5] Figure 8.7. Starting from a racemic mixture, the lipase catalyzes the hydrolysis of the unwanted enantiomer to the carboxylic acid, which spontaneously decarboxylates.

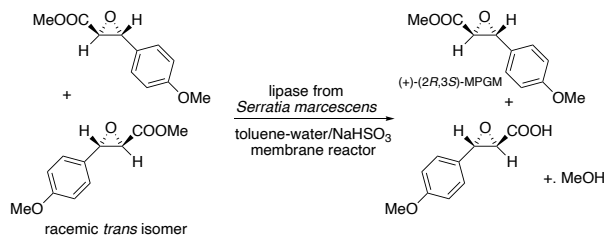


Figure 8.7. Lipase-catalyzed kinetic resolution to manufacture of an enantiopure precursor to diltiazem, a calcium-channel blocker drug. The lipase catalyzes the hydrolysis of the unwanted enantiomer to the carboxylic acid, which spontaneously decarboxylates to the aldehyde. The membrane reactor

In some cases, reversing the reaction reverses its classification as substrate selective or product selective. Reversing the reaction in Figure 8.3 by starting with a racemic mixture of amines and converting them to the ketone would create a case of substrate selectivity. The transaminase would convert the favored (*R*)-amine to the ketone leaving the (*S*)-amine unchanged and an example of a kinetic resolution. The disadvantage of a kinetic resolution is that the maximum yield of each enantiomer is 50%. If the purpose of the reaction is the manufacture of a precursor to make an enantiomerically-pure drug, then only one of the enantiomers will be useful. In contrast, the maximum yield of pure enantiomer in the product selectivity case is 100%.

8.3.1 Measuring selectivity in kinetic resolutions

One way to measure reaction selectivity in a substrate selectivity case is to measure the amounts of substrates consumed or products formed at very low conversion (<5%), eqs. 8.8 or 8.9. At low conversion

$$S_{\text{substrate}} = \frac{(\text{amount of substrate reacted})_f}{(\text{amount of substrate reacted})_s} \text{ at low conversion} \quad (8.8)$$

$$S_{\text{substrate}} = \frac{(\text{amount of product formed})_f}{(\text{amount of product formed})_s} \text{ at low conversion} \quad (8.9)$$

The low conversion is required so that the ratio of the two substrates remains constant over the time of the measurement. The equations 8.8 and 8.9 above also assume equal starting concentrations of the two substrates.

It is often inconvenient to measure the amounts of substrate consumed or product formed at low conversion. In these cases one must use more complex equations. These equations account for the fact that the composition of the substrate changes as the reaction proceeds. The relative amount of the slow-reacting substrate increases as the reaction depletes the fast-reacting substrate, Figure 8.8.

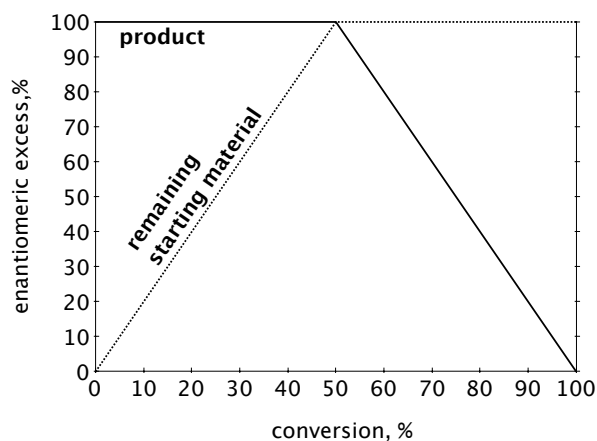


Figure 8.8. Variation of enantiomeric excess for a perfect kinetic resolution as a function of conversion. As the reaction proceeds from 0% to 50% conversion, the remaining starting material is enriched in the slow reacting enantiomer. At 50% conversion, the product consists of the fast-reacting enantiomer, while the substrate consists of the slow reacting enantiomers. The reaction should be stopped at this point. If the reaction is continued, then the enantiomeric excess of the product will decrease as it becomes contaminated with the slow-reacting enantiomer.

The amount of enrichment varies depending on the enantioselectivity of the reaction, Figure 8.9

The equation below^[6] predicts the enantioselectivity, E , from the conversion (c , value between zero and one) and the enantiomeric excess of the product, ee_p , eq. 8.10.

$$E = \frac{\ln[1 - c(1 + ee_p)]}{\ln[1 - c(1 - ee_p)]} \quad (8.10)$$

If you know the value of conversion and the enantiomeric excess of the product, then it is straightforward to calculate the enantioselectivity. For example, if the enantiomeric excess of the product is 90% at 40% conversion, then the enantioselectivity of this reaction is

$$E = \frac{\ln[1 - (0.40)(1 + 0.90)]}{\ln[1 - (0.40)(1 - 0.90)]} = \frac{\ln[(0.24)]}{\ln[(0.96)]} = 35$$

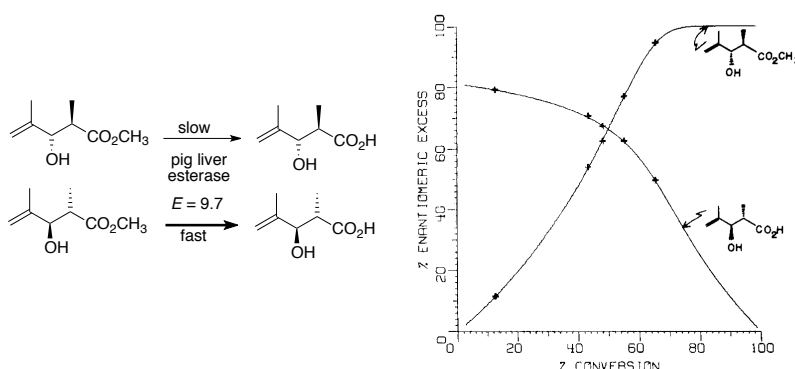


Figure 8.9. An imperfect kinetic resolution. The kinetic resolution of the racemic ester shown catalyzed by pig liver esterase has a moderate enantioselectivity of 9.7. The maximum enantiomeric excess of the product is approximately 80% as low conversion, while the enantiomeric excess of the remaining starting material can be >95% as high conversion.

A web tool encodes eq. 8.10 to simplify this calculation: <http://biocatalysis.uni-graz.at/biocatalysis-tools/enantio>. It also encodes similar equations that can calculate enantioselectivity from ee_s and c or from ee_p and ee_s . Thus, one can calculate the enantioselectivity of a kinetic resolution using any two of the following values: ee_p , ee_s , or c .

However, if you already know the enantioselectivity and would like to calculate the conversion needed to get a certain enantiomeric excess of the product, then you must use an iterative method to solve the equation because you cannot rearrange eq. 8.11 to solve for c . For example, if your reaction has an enantioselectivity of 35 and you would like the enantiomeric excess of the product to be at least 92%, then

$$35 = \frac{\ln[1 - c(1.92)]}{\ln[1 - c(0.08)]} \quad (8.11)$$

Solving this equation iteratively using WolframAlpha or a Python script in the supporting information yields $c = 0.295$. (Less elegantly, one can also try different values of c with $ee_p = 92\%$ in the web tool cited above until the calculated enantioselectivity matches 35.) If you try to solve eq. 8.11 for c when ee_p is a higher value of 95%, you get a negative value for c indicating that it is impossible. An enantioselectivity of 35 is too low to ever yield product with 95% ee.

8.4 Molecular basis of selectivity

Since stereoisomers differ only in shape, the only way to distinguish them is by these differences in shape. Based on many yeast catalyzed reductions of ketones, Prelog proposed a simple model or rule that the favored enantiomer of the product alcohol based

References

1. Coller, B. S. (1985). A new murine monoclonal antibody reports an activation-dependent change in the conformation and/or microenvironment of the platelet glycoprotein IIb/IIIa complex. *J. Clin. Invest.*, 76(1), 101–108. <https://doi.org/10.1172/JCI11931>
2. Novick, S. J., Dellas, N., Garcia, R., Ching, C., Bautista, A., Homan, D., Alvizo, O., Entwistle, D., Kleinbeck, F., Schlama, T., & Ruch, T. (2021). Engineering an amine transaminase for the efficient production of a chiral sacubitril precursor. *ACS Catal.*, 11(6), 3762–3770. <https://doi.org/10.1021/acscatal.0c05450>
3. Savile, C. K., Janey, J. M., Mundorff, E. C., Moore, J. C., Tam, S., Jarvis, W. R., Colbeck, J. C., Krebber, A., Fleitz, F. J., Brands, J., Devine, P. N., Huisman, G. W., & Hughes, G. J. (2010). Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science*, 329(5989), 305–309. <https://doi.org/10.1126/science.1188934>
4. Schechter, I., & Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.*, 27(2), 157–162. <https://doi.org/10.1016/j.bbrc.2012.08.015>
5. Shibatani, T., Omori, K., Akatsuka, H., Kawai, E., & Matsumae, H. (2000). Enzymatic resolution of diltiazem intermediate by *Serratia marcescens* lipase: Molecular mechanism of lipase secretion and its industrial application. *J. Mol. Catal. B Enzym.*, 10(1–3), 141–149. [https://doi.org/10.1016/s1381-1177\(00\)00122-3](https://doi.org/10.1016/s1381-1177(00)00122-3)
6. Chen, C.-S., Fujimoto, Y., Girdaukas, G., & Sih, C. J. (1982). Quantitative analyses of biochemical kinetic resolutions of enantiomers. *J. Am. Chem. Soc.*, 104(25), 7294–7299. <https://doi.org/10.1021/ja00389a064>
7. Prelog, V. (1964). Specification of the stereospecificity of some oxido-reductases by diamond lattice sections. *Pure Appl. Chem.*, 9(1), 119–130. <https://doi.org/10.1351/pac196409010119>
8. Blikstad, C., Dahlström, K. M., Salminen, T. A., & Widersten, M. (2014). Substrate scope and selectivity in offspring to an enzyme subjected to directed evolution. *FEBS J*, 281(10), 2387–2398. <https://doi.org/10.1111/febs.12791>
9. Varadarajan, N., Rodriguez, S., Hwang, B. Y., Georgiou, G., & Iverson, B. L. (2008). Highly active and selective endopeptidases with programmed substrate specificities. *Nat. Chem. Biol.*, 4(5), 290–294. <https://doi.org/10.1038/nchembio.80>
10. Perona, J. J., Hedstrom, L., Rutter, W. J., & Fletterick, R. J. (1995). Structural origins of substrate discrimination in trypsin and chymotrypsin. *Biochemistry*, 34(5), 1489–1499. <https://doi.org/10.1021/bi00005a004>
11. Ekroos, M., & Sjogren, T. (2006). Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc. Natl. Acad. Sci. U. S. A.*, 103(37), 13682–13687. <https://doi.org/10.1073/pnas.0603236103>

12. Demming, R. M., Hammer, S. C., Nestl, B. M., Gergel, S., Fademrecht, S., Pleiss, J., & Hauer, B. (2019). Asymmetric enzymatic hydration of unactivated, aliphatic alkenes. *Angew. Chem. Int. Ed.*, 58(1), 173–177. <https://doi.org/10.1002/anie.201810005>
13. Honaker, M. T., Acchione, M., Sumida, J. P., & Atkins, W. M. (2011). Ensemble perspective for catalytic promiscuity: Calorimetric analysis of the active site conformational landscape of a detoxification enzyme. *J. Biol. Chem.*, 286(49), 42770–42776. <https://doi.org/10.1074/jbc.M111.304386>

Problems

1. Draw an achiral molecule and identify its plane of symmetry. Draw two regioisomers that can result from the electrophilic aromatic substitution of toluene with chlorine. Draw a pair of achiral diastereomers and a pair of chiral diastereomers.

Supporting Information

Code Block S8.1 Python script to calculate the enantioselectivity of a reaction using eq. 8.10.

```
"""
This script calculates the enantioselectivity, E, of a
kinetic resolution from the measured values of percent
conversion and enantiomeric excess of the product in percent.
"""
import math # load math functions needed to calculate logarithms

# enter values for percent conversion and enantiomeric excess of the product
percentconv = raw_input('What is the conversion in percent? (default = 25)') or 25
percenteeep = raw_input('What is the enantiomeric excess of the product in percent? (default = 0)') or 0

# calculate enantioselectivity and round to one digit after decimal
def ecalc(percentconv,percenteeep):
    conv = percentconv/100.0 # convert percent to a fraction between 0 & 1
    eeep = percenteeep/100.0
    E = round((math.log(1-conv*(1+eeep)))/(math.log(1-conv*(1-eeep))),1)
    # math.log(x) calculates the natural logarithm of x
    return E
print 'If the enantiomeric excess of the product is',percenteeep,'\b% ee at',percentconv,'\b%
```