1

Enzymatic C—C Bond Formation by Aldol Reactions

Aldolases

In order to synthesise/metabolise carbohydrates Nature has invented the aldol reaction. Proteins catalysing this reaction between a ketone such as **1** as a nucleophilic donor and an electrophilic aldehyde acceptor **2** leading to a β-hydroxy ketone **3** are called aldolases (Figure 1.1). In general, this reaction is reversible; however, for aldolases, the equilibrium constant can approach $10^4\ M^{-1}$ in favour of the aldol product.

These enzymes are ubiquitous in Nature and have been isolated from mammals, plants, fungi, bacteria, and even thermophylic archaebacteria [1]. Though proteins from different sources are divers with respect to amino acid sequences and global structure they share $(\alpha/\beta)_8$ barrels that contain active sites displaying two modes of activation of the ketone donor: (i) Typ I aldolases employ organocatalysis and (ii) Typ II aldolases use metalorganic activation. Accordingly, aldolases circumvent the use of strong bases – pK_a>18 – that are required for the formation of solvated metal enolates under conventional chemical conditions.

Class I Aldolases

Class I aldolases have been identified more than 80 years ago [2]. Most of these proteins have a molecular weight close to 158 kDa and high sequence homology within one species. The structures of several members of this enzyme family have been determined by X-ray crystallography at high resolution [3] revealing the association of four identical subunits, each containing an active site with a catalytically important lysine. Investigations of aldolases substrate specificity showed a clear preference for DHAP-like nucleophilic ketones, whereas the acceptance of aldehydes is quite broad. However, in all cases the overall stereospecificity of C—C bond formation is the same as explained in the following example.

Fructose 1,6-diphosphate aldolase catalyses the reversible reaction of dihydroxyacetone phosphate **4** (DHAP) with D-glyceraldehyde 3-phosphate **5** (G3P) in favour of the product D-fructose 1,6-diphosphate **6** (FDP) (Figure 1.2).

Bioorganic and Enzymatic Catalysis, First Edition. Wolf-D. Woggon.

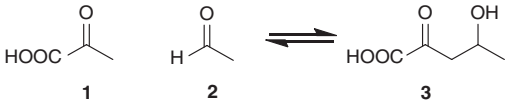

Figure 1.1 The aldol reaction.

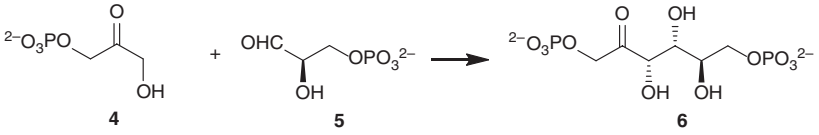

Figure 1.2 Aldolase reaction.

DHAP **4** binds first to the enzyme [4] and reacts with a lysine in the active site to form the Schiff base **7** (Figure 1.3). It is important to note that the pK_a value of the Schiff base-forming lysine must be significantly perturbed in the active site because at physiological pH values lysine is protonated and hence not nucleophilic. For aldolases, two factors adjust the lysin's pK_a value to around pH 7: a hydrophobic, low dielectric environment disfavours charged residues, and a positively charged residue in close proximity further decreases the pK_a of the reactive lysine. A second protonated lysine has been identified indeed in a particular aldolase (DERA) [5].

Evidence for the formation of **7** derives from two experiments (Figure 1.3): (i) on incubation of ^{18}O=C-labelled DHAP the ^{18}O label is released into the medium [6] and (ii) reduction of ^{14}C-labelled **7** yields a catalytically inactive, radioactive enzyme **8** from which the amine **9** can be isolated after hydrolysis of the protein [7]. When

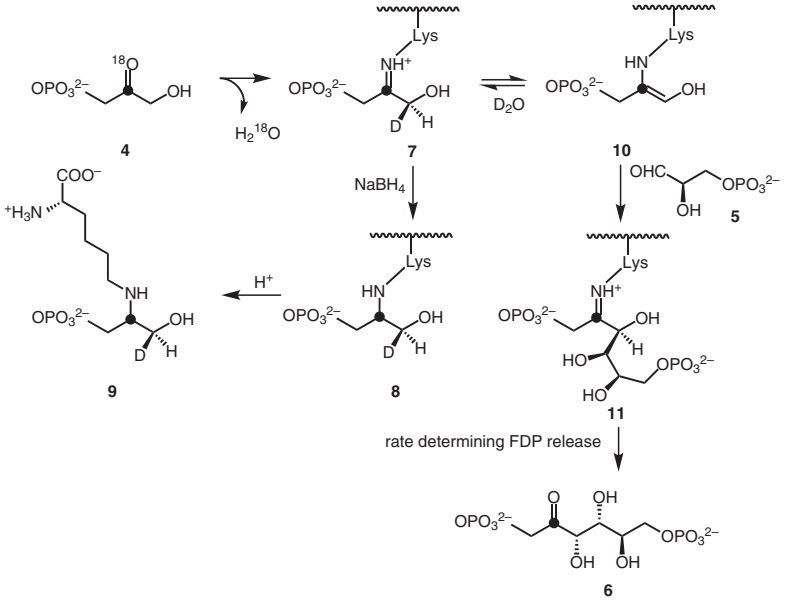

Figure 1.3 Type I aldolase mechanism, • ^{14}C label.

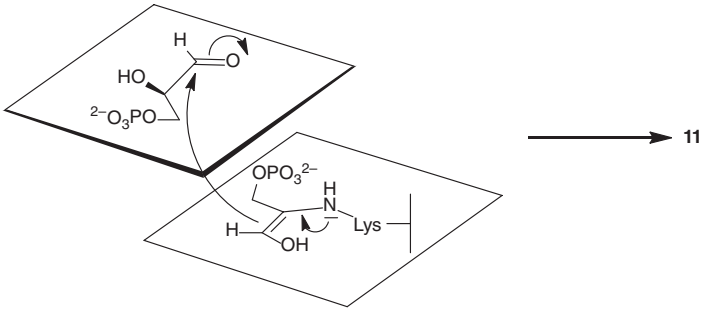

Figure 1.4 Approach of G3P **5** towards the enzyme-bound DAHP **10**.

incubation of **4** is pursued in the absence of **5** in D_2O [8] or tritiated water only, the *pro*-3S proton at C-3 of **7** is exchanged by means of the equilibrium between Schiff base **7** and enamine **10**.

It follows that protonation/deuteration of **10** occurs stereospecific from the *si*-face of the enamine. Interestingly, this is the same face where C—C bond formation occurs. That is the aldehyde **5** approaches the enamine **10** diastereospecifically to give the *S*-configuration at C-3 and the *R*-configuration at C-4 of **11** and **6**, respectively (Figure 1.4).

Due to the stereospecificity of Typ I aldolases and the possibility generating several new stereocentres these enzymes are of synthetic value [9, 10], and many are commercially available and used in industrial processes. Besides the above-mentioned keto aldolases a very interesting aldolase, DERA [5], accepts aldehydes as donors and acceptors for the aldol reaction; even reactions with three to four substrate aldehydes can be performed [11] yielding sugar analogues and key intermediates for the preparation of, e.g. statins and epothilones.

Bioorganic Class I Aldolase Reactions

Since Class I aldolase reactions proceed via enantioface-selective addition of an enamine ketone to an aldehyde, it was obvious that chemists would envisage mimicking this process by preparing chiral enamines *in situ* expecting on addition of an aldehyde chirality transfer to the aldol product. For this purpose, the amino acid proline **12** was the first catalyst of choice due to the presence of a secondary amine for Schiff base formation and a COOH group that could help to organise a stereospecific approach of the acceptor aldehyde/ketone to the enamine by means of H-bonding activation.

The first promising example of a proline-assisted intramolecular aldol reaction was published some 50 years ago [12] using 3 mol% (*S*)-proline **12** to convert the triketone **13** into the ketol **14** and subsequently into the Wieland–Miescher ketone **15** that has been a key intermediate for steroid synthesis at his time (Figure 1.5).

This conversion proved to be extremely efficient generating the diketone **15** in 84% yield and 94% ee. Ever since, the mechanism of this conversion has been a

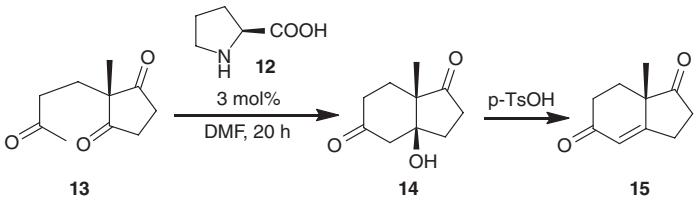

Figure 1.5 Hajos–Perrish–Eder–Sauer–Wiechert (HPESW) reaction.

matter of debate. Analysis of stereoelectronic and steric factors that govern the reaction [13] and DFT calculations [14] led to the conclusion that at least for the HPESW reaction the enamine pathway is the most likely route. This was confirmed by an experiment in the presence of $H_2{}^{18}O$ leading to **^{18}O-14** and **^{18}O-15** labelled at the keto group that formed the enamine/imine (Figure 1.6) [15].

More than 30 years later, stimulated by first results regarding an intermolecular 'direct' aldol reaction [16] catalysed by proline (Figure 1.7), the research field of 'organocatalysis' exploded and hundreds of (*S*)-proline derivatives were tested for catalytic efficiency.

The most effective proline-derived catalysts belong to the class of diaryl prolinols such as **16** that assisted, for example, in the cross-aldol reaction of *o*-chloro benzaldehyde **17** and acetaldehyde **18** to furnish after reduction the diol **19** in 85% yield and 99% ee (Figure 1.8) [17]. In the proposed transition (TS) state **20**, the aldehyde **17** reacts on the more hindered face of the intermediate enamine due to H-bonding between the aldehyde and the OH group of the prolinol subunit.

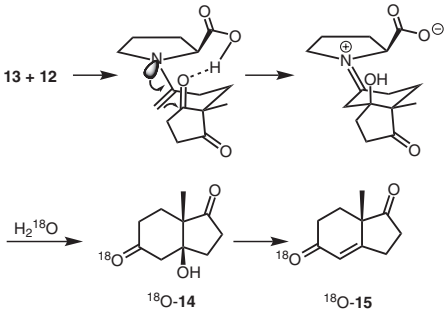

Figure 1.6 Mechanism of the HPESW reaction.

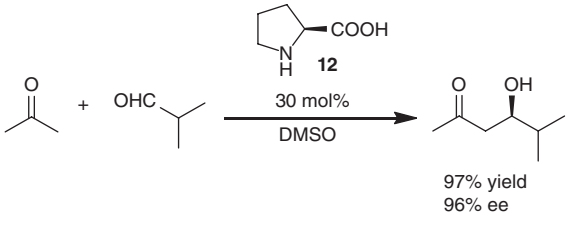

Figure 1.7 (*S*)-Prolin-assisted intermolecular aldol reaction.

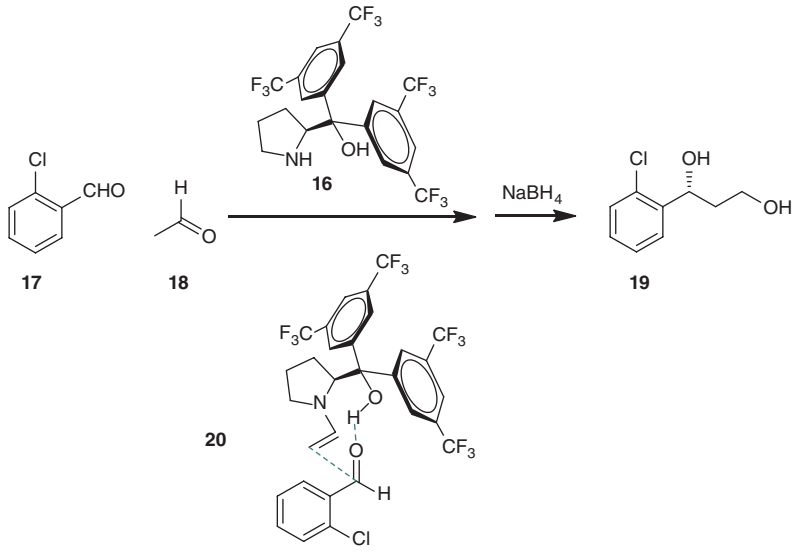

Figure 1.8 (*S*)-Prolin-assisted intermolecular aldol reaction.

The concept of H-bonding-assisted organocatalysis has been exploited using, for example, chiral binaphthyl derivatives **21** and diamino cyclohexanes such as **22**. Both classes of compounds contain two different amine functions of which one forms the enamine and the other being acidic or protonated to provide H-bonding. Using **21** in a cross-aldol reaction of aromatic aldehyde **23** and aldehyde **24** excellent diastereo- and enantioselectivity, see **25**, were observed at relatively low catalyst loading (Figure 1.9) [18].

When the same aromatic aldehyde **23** was reacted with ketone **26** in the presence of diamine **22** the hydroxy ketones **27** were obtained in very good diastereo- and enantioselectivity (Figure 1.10) [19]. However, catalyst loading of ≥10 mol% is required to furnish products in good yield. This unfortunately is a general phenomenon in organocatalysis that catalyst loading is often between 10 and 30 mol% which in principle questions the term catalysis as the turnover is rather low and rarely approaches the corresponding figures of metalorganic catalysis.

Several other very successful organocatalysts such as the prolinol silylether S-**28** [20] and the imidazolidinone **29** [21, 22] omit H-bonding features and rely

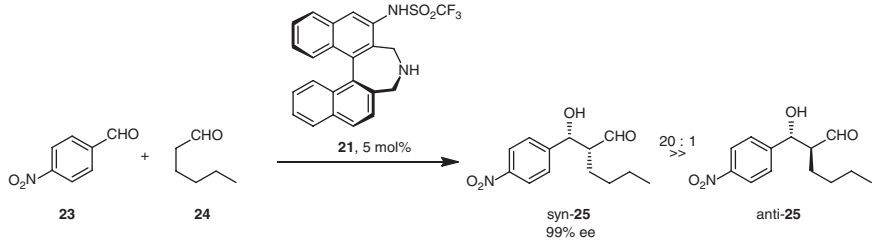

Figure 1.9 Cross-aldol reaction catalysed by sulfonamide **21**.

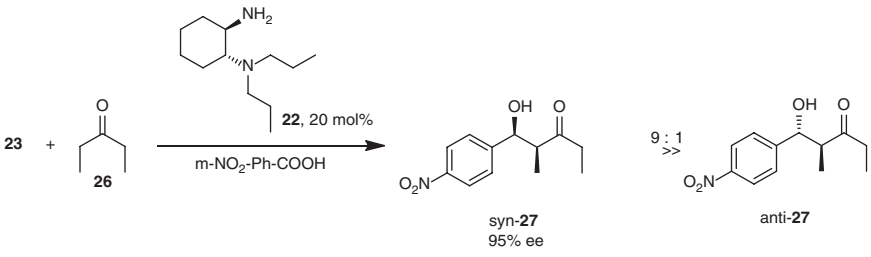

Figure 1.10 Direct aldol reaction catalysed by diamine **22**.

entirely on steric hindrance inducing chirality into aldol products. These compounds have been shown being selective not only for cross-coupling of aldehydes **30** and **31** to **32** (Figure 1.11) but also for a broad spectrum of other reactions involving enamine and dienamine intermediates [20–23].

The prolinol silylether (*R*)-**28** has been used for a quite efficient synthesis of α-tocopherol **33** [23]. The reaction of phytal **34** with (*R*)-**28** gave the dienamine intermediate **35** that subsequently reacts with the aromatic aldehyde **36** yielding the lactol **37** via **38**. The corresponding lactone gave after five steps **33** in 29% overall yield and 93% de, Figure 1.12.

Most of the organocatalytic reactions mentioned above have been done in organic solvents. A more 'green' approach would be if catalysis could be performed in water. This has been realised employing the adamantyl-proline derivative **39** taking advantage of the fact that the adamantyl unit binds to β-cyclodextrin **40** with $K_{ass} = 1.4 \times 10^4\,mol^{-1}$. Hence the catalyst **41** is completely water soluble, and aldol reactions with various aromatic aldehydes and cyclohexanone gave good yields, high *anti/syn* ratios, and ee-values of the dominant *anti* product between 92% and 99%. A few examples are displayed in Figure 1.13 [24].

Though low-molecular-weight organocatalysts have been very successful, **catalytic monoclonal antibodies** proved to be an excellent alternative mimicking aldolase reactions. For the first generation of catalytic antibodies, haptens that resemble the transition state of a reaction were used for immunisation, taking into account

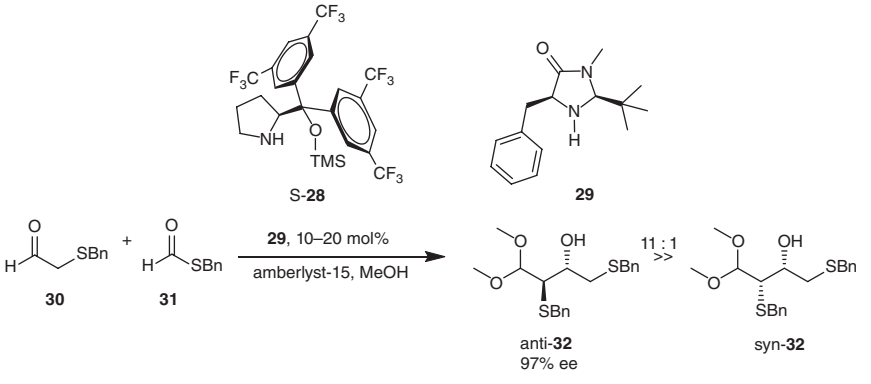

Figure 1.11 Cross-aldol reaction catalysed by imidazolidinone **29**.

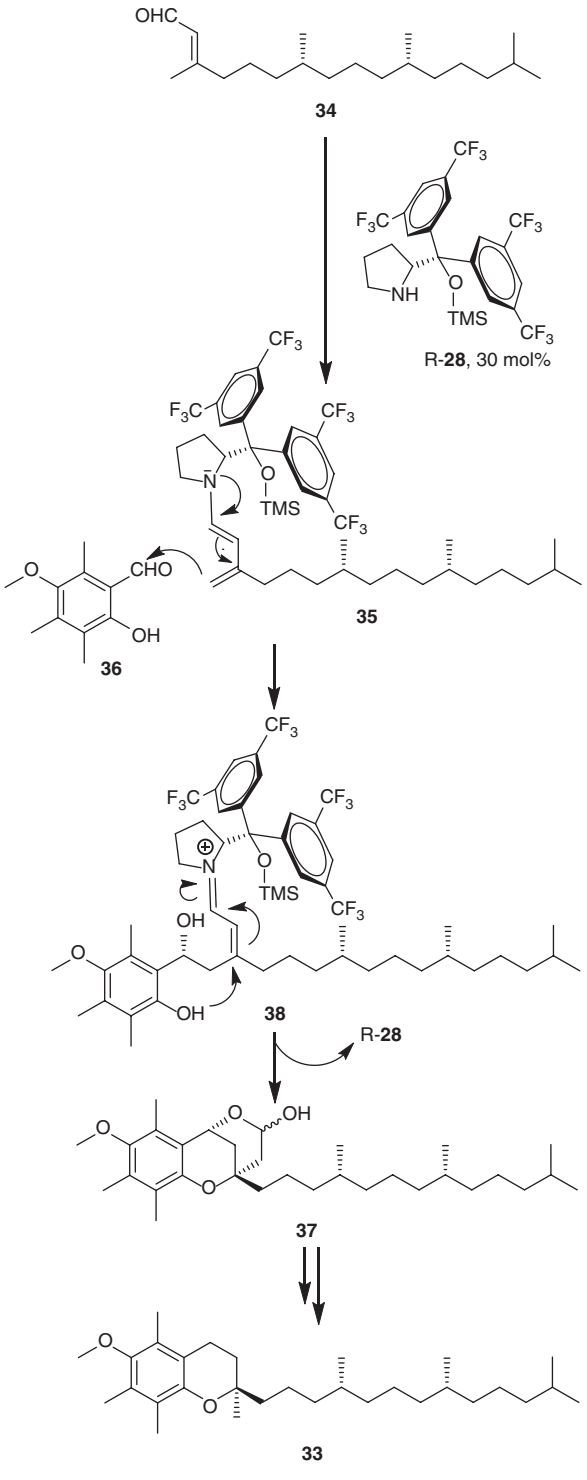

Figure 1.12 Organocatalytic approach to α-tocopherol.

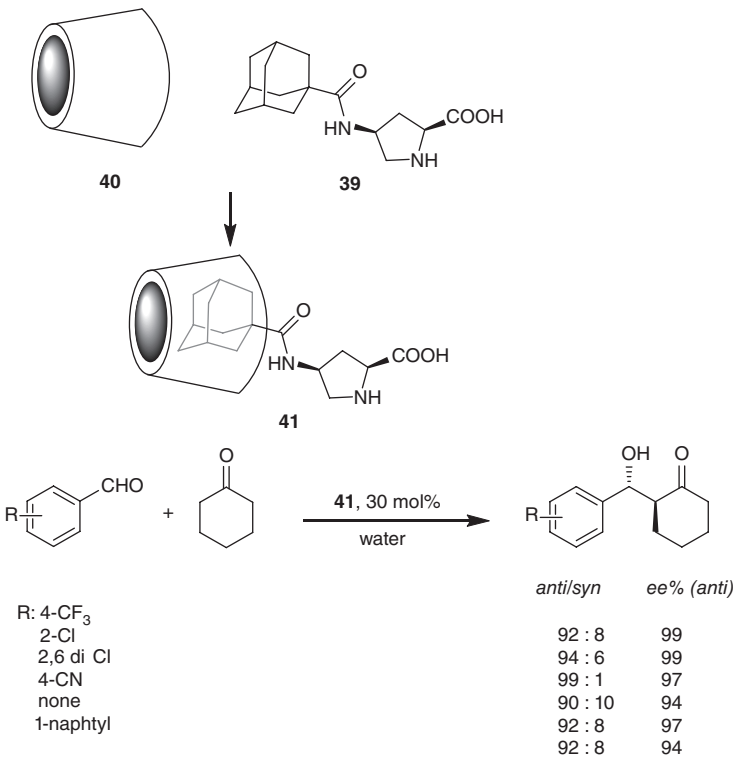

Figure 1.13 Organocatalytic aldol reactions in water.

that enzymes lower the energy of transition states and hence enhance the reaction rate [25]. These haptens were designed to bind to the variable Fab fragment of an antibody by electrostatic interactions and hydrogen bonding. It turned out that monoclonal antibodies generated by transition state analogues (TSA) rarely accomplished rate acceleration comparable with the respective enzymes. This was attributed to the fact that distances between reactands and the polarity in the proposed transition (TS) state are different from the TSA, hence the binding of TS is suboptimal.

To illustrate the distance problem the Claisen rearrangement catalysed by chorismate mutase is shown in Figure 1.14 [26]. The enzyme catalyses the transformation of chorismate **42** to prephenate **43** via TS **44** with rate acceleration of 10^6 over the uncatalysed thermal process [27]. Antibodies were generated against a potent inhibitor of the enzyme, TSA **45**, that catalyse specifically the reaction of (-)-**42** displaying rate acceleration of 10^2–10^4. Calculations revealed that the 'bonds to be created' in TS **44** are significantly longer than the bonds in TSA **45**, suggesting that the FAB binding site is too compact. Further, in order to use the TSA as a hapten, the molecule had to be connected via a linker to the Keyhole Limpet Haemocyanin protein (KLH). This necessary modification also changes the affinity of the binding site for **44**.

Overall, such effects explain, at least in part, the smaller rate enhancement of antibodies against transition states relative to the native enzyme.

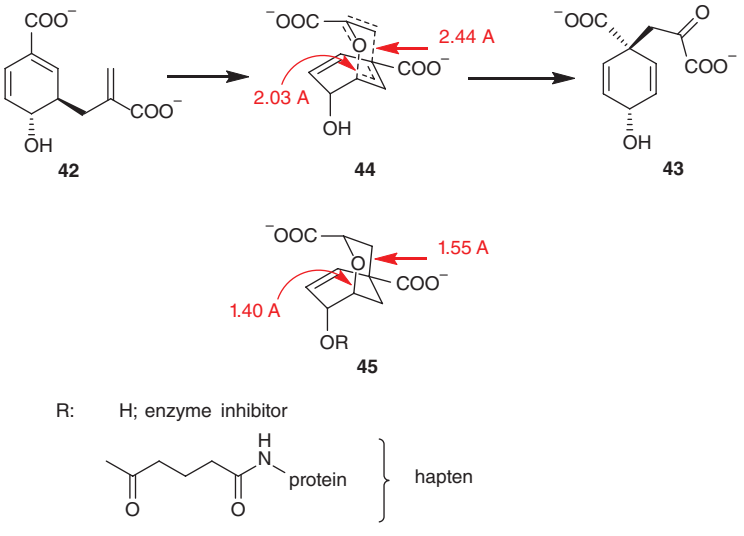

Figure 1.14 Differences between TS **44** and TSA **45**.

In this context a new concept to generate catalytic antibodies termed 'reactive immunisation' was developed [28]. Hence, the *in vivo* selection was not based on tight binding of a 'transition state analogue' but rather on the ability to carry out the aldol reaction. One of the catalytically active mABs was used to bind the diketone **46**. Interestingly the ε-amino group of the mABs lysine was trapped and a stable, covalently bound enaminone **47** was obtained. **47** was spectroscopically identified ($\lambda_{max} = 318$ nm) and the X-ray structure of this Fab-modified monoclonal antibody revealed that the mAB-lysine has reacted with the C-2 keto group of **46** (Figure 1.15) [29].

Further, this mAB-lysine is buried in a hydrophobic cleft adjusting its pK_a and hence making this lysine nucleophilic. In contrast to the rest of the molecules, the enaminone subunit **47** is well ordered in the electron density maps accounting for the broad substrate acceptance by the antibody. Most significant for catalysis is the presence of tyrosine close to the enaminone. The tyrosine acts as a general acid/base catalyst participating in the formation of the enaminone. For the aldol reaction this tyrosine is suggested to form a hydrogen bond to the C=O group of the acceptor aldehyde, see **48**. This hydrogen bond activates the aldehyde, determines the enantioface selectivity, and by stabilising the transition state of the C—C bond-forming step enhances the reaction rate. Accordingly, antibodies were obtained that display catalytic efficiency of $k_{cat}/k_{uncat} \geq 10^5$ for the aldol reaction, which is close to the value of aldolase enzymes but with a broader substrate specificity than natural enzymes [30].

Class II Aldolases

Class II Aldolases are found in yeast, bacteria, fungi, and blue-green algae (*Cyanobacteria*). As they are more stable than Class I aldolases and are absent in

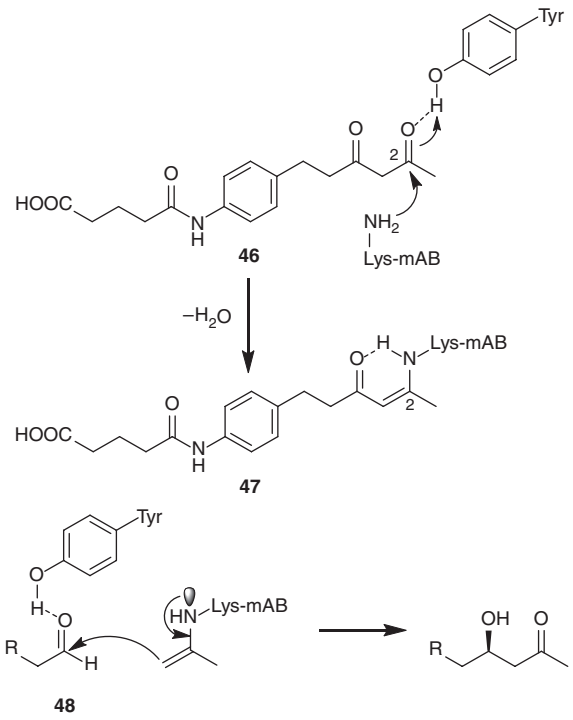

Figure 1.15 In the binding site of the antibody the haptene **46** reacts with the ε-amino group of an available lysine; the role of an adjacent tyrosine is significant for the formation of the adduct **47** as well as for the aldol synthesis of acetone and an aldehyde $R-CH_2-CHO$.

mammals they offer new opportunities for bio-transformation chemistry and as possible targets for anti-bacterial drugs, respectively. Class II aldolases most often contain homodimeric $(\alpha/\beta)_8$ barrels and require for catalysis a Lewis-acidic, divalent metal cation such as Zn^{2+} that can be replaced by Cd^{2+}, Co^{2+}, Fe^{2+}, and Mn^{2+}, obtaining catalytically active enzyme preparations [31, 32]. Hence, these enzymes use the principle of metalorganic catalysis in order to activate substrates. The most detailed structural and mechanistic information is available from investigations of Zn-dependent enzymes, in particular those catalysing the formation/cleavage of FDP **6** (Figure 1.2). The X-ray structure of FDP-aldolase from *E. coli* shows the catalytically active Zn^{2+} coordinating to three histidines and one glutamate in a distorted tetrahedral geometry, see **49**, Figure 1.16 [33]. The glutamate can be replaced by the transition state analogue phosphoglycolohydroxamate **50** (PGH), and the resulting structure **51** demonstrates the coordination of both the OH group and O^- of **50** to Zn [34].

This result indicates that DHAP **4** binds to Zn as the enolate, see **52** (Figure 1.17), that is approached by the aldehyde G3P **5** yielding the enzyme-product complex **53** finally releasing the product FDP **6**. This stereospecific addition of **5** is governed mainly by strong Coulomb interaction of the phosphate group of **5** to an arginine of the protein and an aspartate activating the C=O group of the aldehyde; hence, only the *si*-face of the *syn*-enolate, see **52**, can be approached.

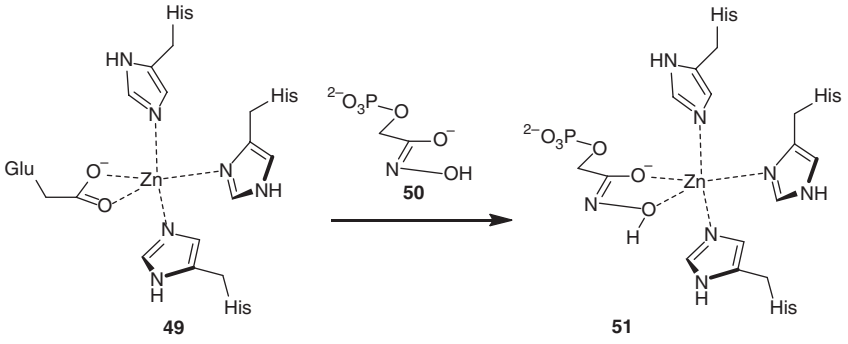

Figure 1.16 Binding of the transition state analogue **50** in the active site of FDP aldolase from *E. coli*.

Besides the substrate-binding Zn two other cations, another Zn^{2+} and one monovalent cation, are important for catalysis. The presence of K^+ or Na^+ has been shown enhancing the rate of the reaction, and the X-ray structures suggest that these monovalent cations interact with the phosphate group of DAHP aligning the substrate for catalysis. The second Zn^{2+} adjacent to the catalytic zinc site is created near the interface of the homodimeric protein, and it is believed to order and stabilise the catalytic site [34].

Class II Aldolase Mimics

The first approaches to mimic enantioselective aldolase reactions focused on the addition of activated electrophiles, for example electron-rich O-silyl enol ethers **54**, to conjugated aldehydes **55** catalysed by chiral Lewis acid complexes such as **56** [35].

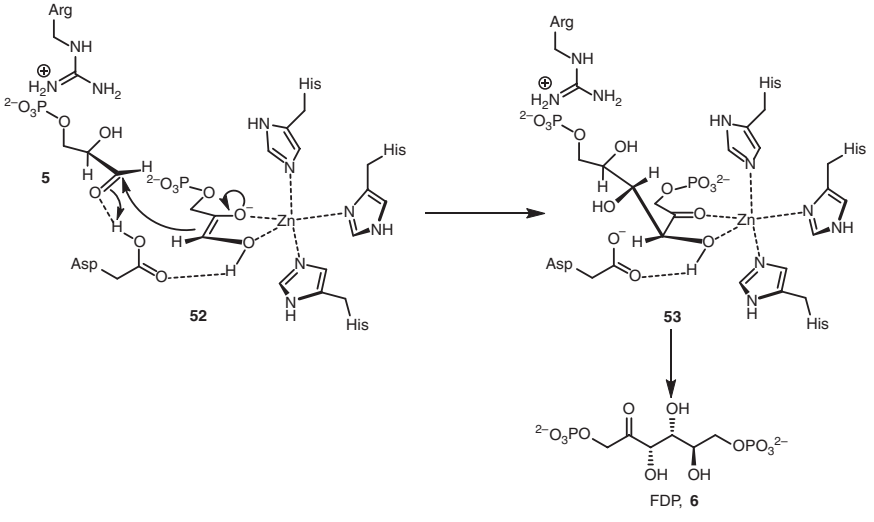

Figure 1.17 Addition of G3P **5** to the Zn-activated DHAP **52**.

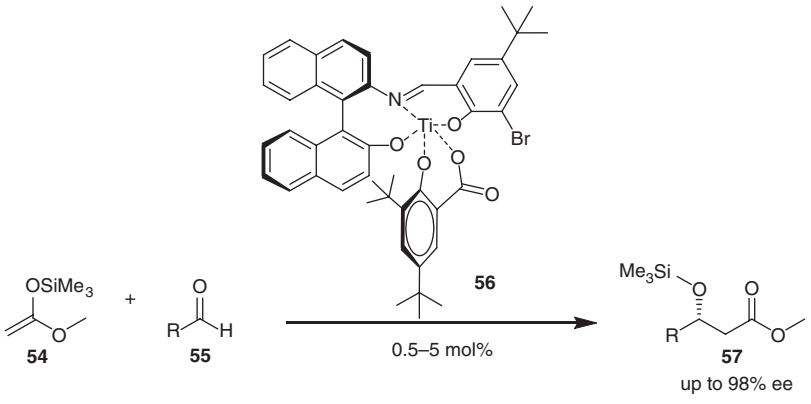

Figure 1.18 Enantioselective aldol reaction with silyl ketene acetals.

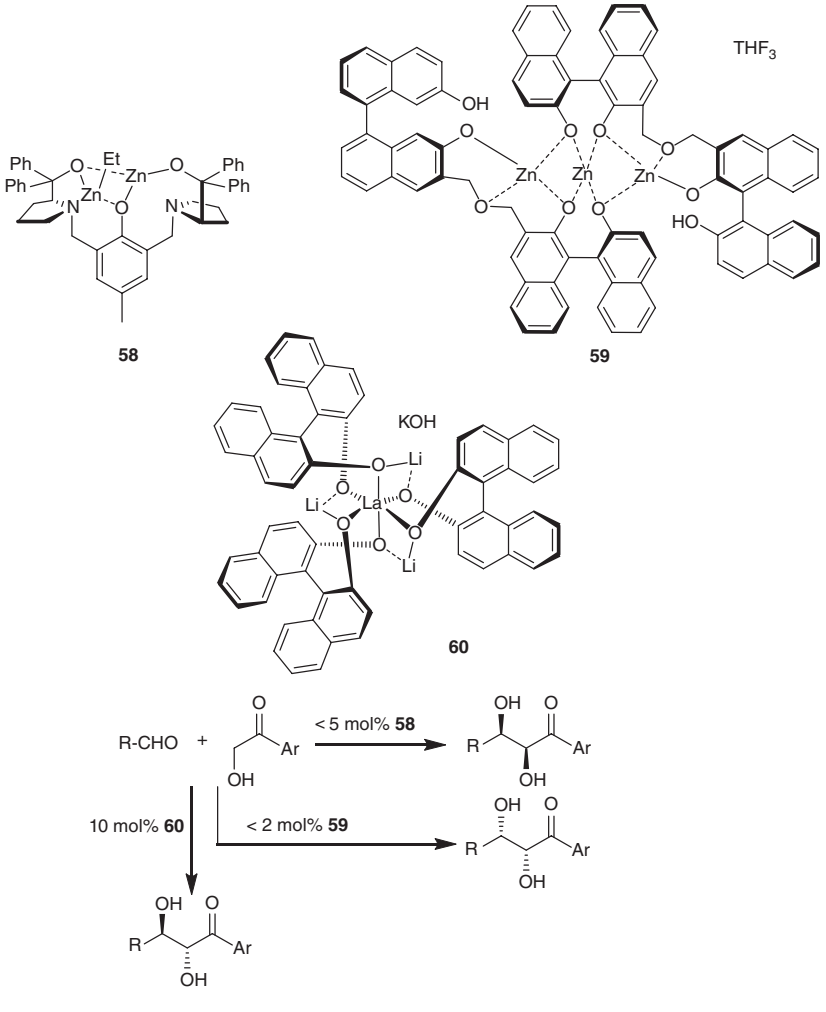

Figure 1.19 Metal complexes for enantio- and diastereoselective direct aldol reactions.

This methodology gave β-hydroxyesters **57** in excellent yields and enantiomeric excess (Figure 1.18) employing very low catalyst loading.

Developing catalysts for the more natural reaction, i.e. the direct asymmetric aldol reaction, has been the aim of many research groups [36]. Most significant are the results with bimetallic **58** [37], trimetallic **59** [38], and hetero bimetallic catalysts **60** (Figure 1.19) [39]. These complexes are very selective regarding the preference for one diastereoisomeric product with ee-values >95%; see, for example, the reactions of α-hydroxy ketones with aldehydes.

References

1 Gefflaut, T., Blonski, C., Perie, J., and Willson, M. (1995). *Prog. Biophys. Mol. Biol.* 63: 301. Lorentzen, E., Siebers, B., Hensel, R., Pohl, E. (2004) *Biochem. Soc. Trans.* 32: 259.

2 Meyerhof, O. and Lohmann, K. (1934). *Biochem. Z.* 271: 89.

3 Fothergill-Gilmore, L.A. and Michels, P.A.M. (1993). *Prog. Biophys. Mol. Biol.* 59: 105.

4 Rose, I.A., O'Connell, E.L., and A.H. Mehler (1965). *J. Biolumin. Chemilumin.* 240: 1758.

5 Heine, A., Luz, J.G., Wong, C.-H., and I.A. Wilson (2004). *J. Mol. Biol.* 343: 1019.

6 Model, P., Ponticarvo, L., and Rittenberg, D. (1968). *Biochemistry* 7: 1339.

7 Grazi, E., Cheng, T., and Horecker, B.L. (1962). *Biochem. Biophys. Res. Commun.* 7: 250.

8 Rutter, W.J. and Ling, K.H. (1958). *Biochem. Biophys. Acta* 30: 71.

9 Müller, M. (2012). *Adv. Synth. Catal.* 354: 3161.

10 Fessner, W.-D. (2008). Aldolases: enzymes for making and breaking C-C bonds. In: *Asymmetric Organic Synthesis with Enzymes* (ed. V. Gotor, I. Alfonso, and E. Garca-Urdiales), 275. Weinheim: Wiley-VCH Verlag.

11 Wolberg, M., Dassen, B.H.N., Schürmann, M. et al. (2008). *Adv. Synth. Catal.* 350: 1751. Whalen, L.J. and Wong, C.-H. (2006) *Aldrichim. Acta* 39: 63.

12 Hajos, Z.G. and Parrish, D.R. (1974). *J. Organomet. Chem.* 39: 1615. Eder, U., Sauer, G., Wiechert, R. (1971) *Angew. Chem. Int. Ed. Engl.* 10: 496.

13 Seebach, D., Beck, A.K., Limbach, M. et al. (2007). *Helv. Chim. Acta* 90: 425.

14 Hoang, L., Bahnmanyar, S., Houk, K.N., and List, B. (2003). *J. Am. Chem. Soc.* 125: 16.

15 List, B., Hoang, L., and Martin, H.J. (2004). *Proc. Nat. Acad. Sci. U.S.A.* 101: 5839.

16 List, B., Lerner, R.A., and Barbas, C.F. III (2000). *J. Am. Chem. Soc.* 122: 2395.

17 Hayashi, Y., Itoh, T., Aratake, S., and Ishikawa, H. (2008). *Angew. Chem. Int. Ed.* 47: 2082.

18 Kano, T., Yamaguchi, Y., Tanaka, Y., and Maruoka, K. (2007). *Angew. Chem. Int. Ed.* 46: 1738.

19 Luo, S., Xu, H., Li, J. et al. (2007). *J. Am. Chem. Soc.* 129: 3074.

20 Jensen, K.L., Dickmeiss, G., Jiang, H. et al. (2012). *Acc. Chem. Res.* 45: 248.

21 Mangion, I.K., Northrup, A.B., and MacMillan, D.W.C. (2004). *Angew. Chem. Int. Ed.* 43: 6722.

22 Lelais, G. and MacMillan, D.W.C. (2006). *Aldrichim. Acta* 39: 79.

23 Liu, K., Chougnet, A., and Woggon, W.-D. (2008). *Angew. Chem. Int. Ed.* 47: 5827.
24 Liu, K., Häussinger, D., and Woggon, W.-D. (2007). *Synlett* 14: 2298.
25 Lerner, R.A., Benkovic, S.J., and Schultz, P.G. (1991). *Science* 252: 659.
26 Barbany, M., Gutiérrez-de-Teràn, H., Sanz, F. et al. (2003). *ChemBioChem* 4: 277.
27 Hilvert, D., Carpenter, S.H., Nared, K.D., and Auditor, M.-T.M. (1988). *Proc. Nat. Acad. Sci. U.S.A.* 85: 4953.
28 Tanaka, F. and C.F. (2005). Barbas III. In: *Catalytic Antibodies* (ed. E. Keinan), 304. Weinheim: Wiley-VCH Verlag.
29 Zhu, X., Tanaka, F., Lerner, R.A. et al. (2009). *J. Am. Chem. Soc.* 131: 18206.
30 Zhu, X., Tanaka, F., Hu, Y. et al. (2004). *J. Mol. Biol.* 343: 1269.
31 Hall, D.R., Kemp, L.E., Leonard, G.A. et al. (2003). *Acta Cryst.* D59: 611.
32 Hill, H.A.O., Lobb, R.R., Sharp, S.L. et al. (1976). *Biochem. J* 153: 551.
33 Cooper, S.J., Leonard, G.A., McSweeney, S.M. et al. (1996). *Structure* 4: 1303.
34 Hall, D.R., Leonard, G.A., Reed, C.D. et al. (1999). *J. Mol. Biol.* 287: 383.
35 Carreira, E.M., Singer, R.A., and Lee, W. (1994). *J. Am. Chem. Soc.* 116: 8837.
36 Shibasaki, M., Matsunaga, S., and Kumagai, N. (2004). *Modern Aldol Reactions. Vol. 2: Metal Catalysis* (ed. R. Mahrwald), 197. Weinheim: Wiley-VCH Verlag.
37 Trost, B.M. and Ito, H. (2000). *J. Am. Chem. Soc.* 122: 12003.
38 Kumagai, N., Matsunaga, S., Kinoshita, T. et al. (2003). *J. Am. Chem. Soc.* 125: 2169.
39 Yoshikawa, N., Kumagai, N., Matsunaga, S. et al. (2001). *J. Am. Chem. Soc.* 123: 2466.