Efficient asymmetric organocatalytic formation of erythrose and threose under aqueous conditions†

Laurence Burroughs,[a] Matthew E. Vale,[ab] James A. R. Gilks,[b] Henrietta Forintos,[b] Christopher J. Hayes*[b] and Paul A. Clarke*[ab]

Received 26th March 2010, Accepted 5th May 2010
First published as an Advance Article on the web 20th May 2010
DOI: 10.1039/c0cc00613k

Esters of proteinogenic amino acids efficiently catalyse the formation of erythrose and threose under aqueous conditions in the highest yields and enantioselectivities yet reported. Remarkably while esters of (L)-proline yield (L)-carbohydrates, esters of (L)-leucine and (L)-alanine generate (D)-carbohydrates, offering the potential to account for the prebiotic link between natural (L)-amino acids and natural (D)-sugars.

It is thought that Life on Earth began more than 3.5 billion years ago. This means that all of the chemical building blocks of primitive Life must have been present on Earth at that time. These materials could have been already present on Earth as a result of its formation or they could have been delivered to the surface as a consequence of impactors (*i.e.* meteorites and asteroids).[1–4] It has been postulated that the first carbohydrates may have been formed in an amino acid catalysed aldol reaction of glycolaldehyde. However, this process is very inefficient in terms of both the yields and enantiopurity of the products. In a recent study Pizzarello and Weber showed that the non-proteinogenic amino acid (*R*)-isovaline (>95% ee) catalysed an aldol dimerisation of glycolaldehyde in water to produce (L)-threose in low % enantiomeric excess (10% ee).[5] Unfortunately the yield of threose produced in this transformation was not reported. Additionally, Darbre *et al*. showed that zinc–proline complexes can catalyse the formation of simple sugars in water.[6] However, no enantioselectivities were reported and as with the Pizzarello and Weber studies the yields were very low.

Recent elegant studies of List, MacMillan, and Barbas have shown that highly asymmetric aldol dimerisations of protected glycolaldehyes can be achieved in organic solvents using proline (100% ee) as the catalyst.[7–13] However, such amino acid catalysed aldol reactions are nowhere near as successful in water, which must be viewed as the solvent for any prebiotic synthesis of carbohydrates.[11,12] We have independently confirmed that both enantiomers of proline are incapable of catalysing the dimerisation of glycolaldehyde, or its triisopropylsilyl (TIPS) protected derivative, in water. Fascinated by these observations, we set out to develop a simple, catalytic, high yielding and potentially prebiotic aldol dimerisation using water as the reaction solvent.

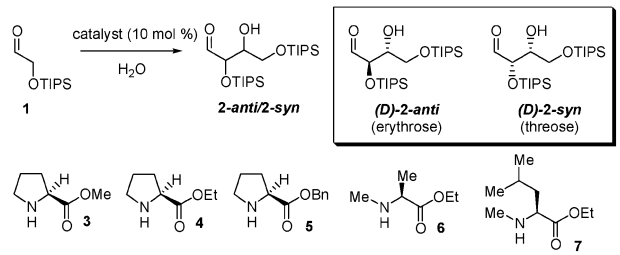
There has been a recent flurry of activity in the area of organocatalysed aldol reactions in water,[14–25] and this has led to some very interesting debates.[21,22] Janda and Dickerson have recently shown that nicotine can act as an organocatalyst for aldol reactions under aqueous conditions, but so far only racemic studies have been described.[23] Further work by Janda *et al*. has shown that C2-aromatic substituted pyrrolidines also serve as aldol-catalysts in water, and that the best examples have electron-withdrawing substituents on the aromatic portion of the molecule.[24] As it is known that proline is a good catalyst for the desired aldol dimerisation in organic solvents, we wondered if simple changes could be made to this amino acid to render it an effective catalyst in water. It is known that the carboxylic acid group of proline plays a key hydrogen-bonding role during the catalytic aldol reactions in organic solvents,[7] and these interactions will probably be lost in aqueous conditions, which almost certainly accounts for proline's lack of catalytic activity in water. We postulated that simple esters of proline could act as aldol catalysts in water, as the resulting pyrrolidine would bear an electron withdrawing group at C2, similar to the effective aromatic substituted catalysts of Janda.

Due to the difficulties inherent in monitoring and analysing the direct aqueous dimerisation of glycolaldehyde, we initially decided to study the dimerisation of the TIPS-protected glycolaldehyde **1**. We reasoned that as this system was known from the studies of MacMillan this would enable us to more easily develop conditions for aldol dimerisation and analysis of the products. Pleasingly, we found that simple esters (methyl, ethyl, benzyl) of (L)-proline **3**, **4**, **5** mediated the dimerisation of TIPS-protected glycolaldehyde **1** in water to afford the desired protected threose and erythrose products **2** (Scheme 1) in good isolated yield with moderate diastereo- and enantioselectivity (see Table 1, entries 1–3). Although the enantioselectivities are moderate these are the highest values yet obtained for the amino acid catalysed formation of threose and erythrose derivatives **2**-*syn*/**2**-*anti* in water.

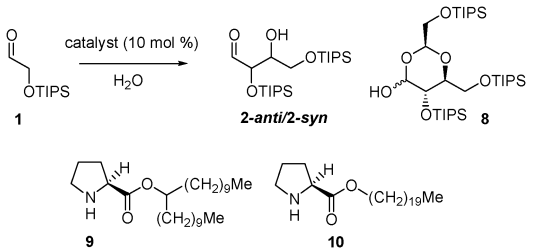
Entry 4, Table 1 shows that the reaction time can be significantly shortened from 48 hours to 5 hours without significantly affecting either isolated yield or stereoselectivity, and as a result all other reactions were performed for 5 hours. We next widened our study to examine the use of simple esters of other α-amino acids, and to this end we screened the ethyl esters of (L)-*N*-methyl alanine **6** and (L)-*N*-methyl leucine **7** as catalysts. Once again we observed the formation of

[a] *Department of Chemistry, University of York Heslington, York, UK YO10 5DD. E-mail: pac507@york.ac.uk; Fax: +44 (0)1904 432516; Tel: +44 (0)1904 432614*
[b] *School of Chemistry, University of Nottingham, University Park, Nottingham, Notts, UK NG7 2RD. E-mail: chris.hayes@nottingham.ac.uk; Fax: +44 (0)115 9513045; Tel: +44 (0)115 9513564*

† Electronic supplementary information (ESI) available: Experimental procedures and compound characterisation data. See DOI: 10.1039/c0cc00613k



Scheme 1 Formation of protected tetroses in water.



Scheme 2 Formation of protected tetroses in water and pH = 7 buffer.

tetrose products, but this time we observed the formation of (D)-threose and (D)-erythrose *ent*-**2**-*anti* as the major enantiomers in good yield.

Although high yields were generally obtained after 5 hours using either the proline, alanine or leucine (*e.g.* entries 4–6) derived catalysts, the % enantiomeric excesses observed were only modest (*i.e.* 22% ee, entry 2, Table 1). Janda *et al.* have suggested that non-selective, general base catalysis of the aldol reaction may operate in situations where amine-derived catalysts are used in aqueous reaction environments,[21] so we next decided to explore this possibility by examining the influence of pH on the course of the reaction. The proline benzyl ester **5** (entry 4, Table 1) mediated aldol dimerisation of TIPS-protected glycolaldehyde **1** in water was chosen as a representative example, and we monitored the pH of the reaction mixture at various time points over a 48 h time period. From this pH profile (see ESI†) it can be clearly seen that the pH of the reaction mixture is basic (pH = 8.5) at the start of the reaction and that it increases to a maximum (pH = 8.8) after approximately 5 hours. After this time the pH drops steadily to a final pH of 7.8 after 48 hours. These observations clearly show that a general base mediated aldol reaction could be operating under these reaction conditions, so we next decided to perform the same aldol dimerisation but this time buffered (KH_2PO_4/NaOH) to pH = 7 (Scheme 2, Table 2).

Under these buffered conditions we were pleased to find that we still obtained a good yield of the desired tetrose aldol products **2** (70%, 1.5 : 1 *anti* : *syn*), but more importantly the % enantiomeric excess of the major product had risen from 15% to 47% ee. We believe that the improved % enantiomeric excess is due to the operation of an enamine-based dimerisation mechanism, with minimal involvement of a non-selective general base-mediated mechanism. We next examined the use of *N*-methyl (L)-leucine ethyl ester **7** under buffered conditions (pH = 7) and we were pleased to see that a good yield (79%; 1.5 : 1 *anti* : *syn*) and the highest % enantiomeric excess yet (57% ee for erythrose) were obtained. As observed previously for the unbuffered reactions, the major product obtained using the leucine-derived catalyst was (D)-erythrose *ent*-**2**-*anti*, thus demonstrating that the change in absolute configuration is catalyst dependent and is not affected by the pH change.

In an attempt to maximise the local hydrophobic environment imposed by the catalyst,[15] we next prepared two long-chain aliphatic ester derivatives of the proline catalyst and tested them in the aldol dimerisation reaction under buffered and unbuffered conditions. The branched alkyl ester (heneicosyl) catalyst **9** gave a moderate yield (49%) and low % enantiomeric excess (10% ee) under the standard unbuffered conditions, but the *anti* : *syn* ratio (2 : 1) was slightly improved when compared to the previously used catalysts. Under the

Table 1 The amino ester catalysed formation of protected erythrose in water

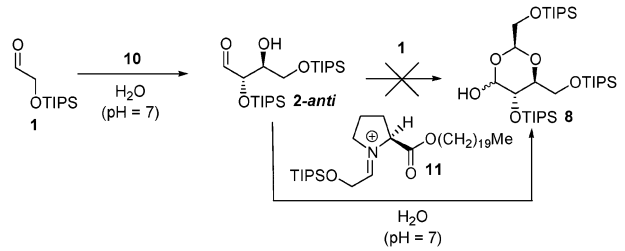
Entry	Catalyst	Major product[c]	Combined yield (%)	Ratio (*anti* : *syn*)	% ee *anti*[d]
1[a]	**3**	(L)-**2**-*anti*	68	4.5 : 1	15
2[a]	**4**	(L)-**2**-*anti*	57	4.0 : 1	22
3[a]	**5**	(L)-**2**-*anti*	80	1.5 : 1	15
4[b]	**5**	(L)-**2**-*anti*	77	1.5 : 1	18
5[b]	**6**	(D)-**2**-*anti*	70	3.0 : 1	7
6[b]	**7**	(D)-**2**-*anti*	80	1.5 : 1	17

[a] Reaction time 48 h. [b] Reaction time 5 h. [c] Major enantiomer determined by correlation to the work of Northrup and MacMillan (ref. 12). [d] See ESI.†

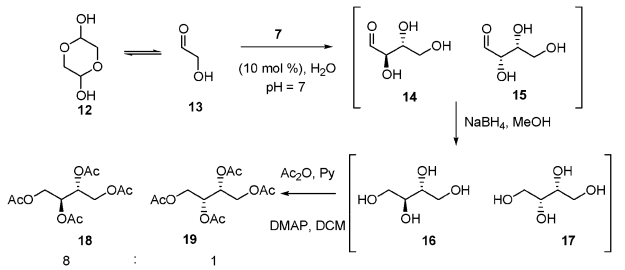
Table 2 The amino ester catalysed formation of protected erythrose in water and pH = 7 buffer

Entry	Catalyst	Major product[c]	Combined yield (%)	Ratio (*anti* : *syn*)	% ee *anti*[d]
1[a]	**5**	(L)-**2**-*anti*	70	1.5 : 1	47
2[a]	**7**	(D)-**2**-*anti*	79	1.5 : 1	57
3[b]	**9**	(L)-**2**-*anti*	49	2.0 : 1	10
4[b]	**9**	(L)-**2**-*anti*	52	5.5 : 1	46
5[b]	**10**	**8**	40	N/A	17
6[b]	**10**	**8**	32	N/A	23

[a] pH 7 buffer. [b] Unbuffered. [c] Major enantiomer determined by correlation to the work of Northrup and MacMillan (ref. 12). [d] See ESI.†



Scheme 3 Amino ester catalysed formation of a cyclic acetal.



Scheme 4 Amino ester catalysed formation of (D)-threose.

pH = 7 buffered conditions, the % enantiomeric excess of the (L)-erythrose product **2**-*anti* was dramatically increased (46% ee) and the *anti* : *syn* ratio was improved to 5.5 : 1. Although the isolated yield was still moderate (52%) the mass balance of the reaction was unreacted starting aldehyde.

In contrast to the results observed previously, only a small amount of the *anti*-aldol dimer was produced (10–11%) when using the straight chain (icosyl) ester catalyst **10** under either the buffered or unbuffered conditions. The major product produced in both cases (entries 5 and 6, Table 2) was the unusual acetal product **8** which is effectively a trimer of the aldehyde starting material. Although the formation of this product was unexpected, especially at pH = 7, its production from the initially formed *anti*-aldol dimer **2**-*anti* can obviously be explained by invoking a simple acetalisation process (Scheme 3).

As we do not see the acetal product **8** in any of the other aldol dimerisations described above, we believe that the acetal **8** is formed from the *anti*-aldol dimer **2**-*anti* by reaction with the catalyst-derived iminium ion **11** rather than by reaction directly with TIPS-glycolaldehyde **1** present in the reaction mixture. The reason for why the long-chain isocosyl ester catalyst **10** prefers this reaction manifold rather than stopping at the aldol dimer **2**-*syn*/**2**-*anti* is not yet clear, but this result is reproducible and is not pH dependent.

Having demonstrated that amino acid-derived ester catalysts are capable of performing aldol dimerisations of TIPS-glycolaldehyde **1** in water, we next examined the aldol dimerisation of glycolaldehyde **13** itself under our optimised conditions. The *N*-methyl leucine ethyl ester catalyst **7** was chosen for this purpose as it provided the best compromise between yield and % enantiomeric excess of the tetrose products under neutral reaction conditions. Thus, an aqueous solution of glycolaldehyde dimer **12** was buffered to pH = 7 as previously described, and *N*-methyl leucine ethyl ester **7** (10 mol%) was added and the mixture was stirred vigorously for 5 hours (Scheme 4).

Due to the high polarity and water solubility of the unprotected tetrose products (erythrose **14** and threose **15**), we decided to reduce the crude reaction mixture ($NaBH_4$, MeOH) to afford erythritol **16** and threitol **17**, which were then acetylated with excess acetic anhydride to afford the corresponding tetra-acetylated products **18** and **19** in a 8 : 1 ratio.[16] The *meso* compound **18** corresponds to the *anti*-aldol product (erythrose **14**) and the C_2-symmetric compound **19** corresponds to the *syn*-aldol product (threose **15**), and pleasingly the threose derived tetra-acetate **19** was produced in 68% ee with (D)-threitol predominating. This is the highest asymmetric induction reported for the direct aldol dimerisation of unprotected glycolaldehyde **13**, in either aqueous or organic solvents, and represents a dramatic improvement over the 10% ee originally reported in the work of Pizzarello and Weber, while additionally offering the potential to account for the link between (L)-amino acids and (D)-carbohydrates.

Notes and references

1 A. Jorissen and C. Cerf, *Origins Life Evol. Biosphere*, 2002, **32**, 129–142.
2 R. Root-Bernstein, *BioEssays*, 2007, **29**, 689–698.
3 Z. Martins, O. Botta, M. L. Fogel, M. A. Sephton, D. P. Glavin, J. S. Watson, J. P. Dworkin, A. W. Schwartz and P. Ehrenfreund, *Earth Planet. Sci. Lett.*, 2008, **270**, 130–136.
4 S. L. Miller and H. C. Urey, *Science*, 1959, **130**, 245–251.
5 S. Pizzarello and A. L. Weber, *Science*, 2004, **303**, 1151.
6 J. Kofoed, J.-L. Reymond and T. Darbre, *Org. Biomol. Chem.*, 2005, **3**, 1850.
7 B. List, R. A. Lerner and C. F. Barbas III, *J. Am. Chem. Soc.*, 2000, **122**, 2395–2396.
8 K. Sakthivel, W. Notz, T. Bui and C. F. Barbas III, *J. Am. Chem. Soc.*, 2001, **123**, 5260–5267.
9 B. List, L. Hoang and H. J. Martin, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 5839–5842.
10 B. List, *Tetrahedron*, 2002, **58**, 5573–5590.
11 A. B. Northrup, I. K. Mangion, F. Hettche and D. W. C. MacMillan, *Angew. Chem., Int. Ed.*, 2004, **43**, 2152–2154.
12 A. B. Northrup and D. W. C. MacMillan, *Science*, 2004, **305**, 1752–1755.
13 A. Córdova, M. Engqvist, I. Ibrahem, J. Casas and H. Sundén, *Chem. Commun.*, 2005, 2047–2049.
14 A. Córdova, W. Notz and C. F. Barbas III, *Chem. Commun.*, 2002, 3024–3025.
15 N. Mase, Y. Nakai, N. Ohara, H. Yoda, K. Takabe, F. Tanaka and C. F. Barbas III, *J. Am. Chem. Soc.*, 2006, **128**, 734–735.
16 A. Córdova, I. Ibrahem, J. Casas, H. Sundén, M. Engqvist and E. Reyes, *Chem.–Eur. J.*, 2005, **11**, 4772–4784.
17 J. Mlynarski and J. Paradowska, *Chem. Soc. Rev.*, 2008, **37**, 1502–1511.
18 F. Rodríguez-Llansola, J. F. Miravet and B. Escuder, *Chem. Commun.*, 2009, 7303–7305.
19 M. De Nisco, S. Pedatella, H. Ullah, J. H. Zaidi, D. Naviglio, Ö. Özdamar and R. Caputo, *J. Org. Chem.*, 2009, **74**, 9562–9565.
20 N. Mase, N. Noshiro, A. Mokuya and K. Takabe, *Adv. Synth. Catal.*, 2009, **351**, 2791–2796.
21 A. P. Brogan, T. J. Dickerson and K. D. Janda, *Angew. Chem., Int. Ed.*, 2006, **45**, 8100–8102.
22 Y. Hayashi, *Angew. Chem., Int. Ed.*, 2006, **45**, 8103–8104.
23 T. J. Dickerson and K. D. Janda, *J. Am. Chem. Soc.*, 2002, **124**, 3220–3221.
24 C. J. Rogers, T. J. Dickerson, A. P. Brogan and K. D. Janda, *J. Org. Chem.*, 2005, **70**, 3705–3708.
25 S. Aratake, T. Itoh, T. Okano, T. Usui, M. Shoji and Y. Hayashi, *Chem. Commun.*, 2007, 2524.

