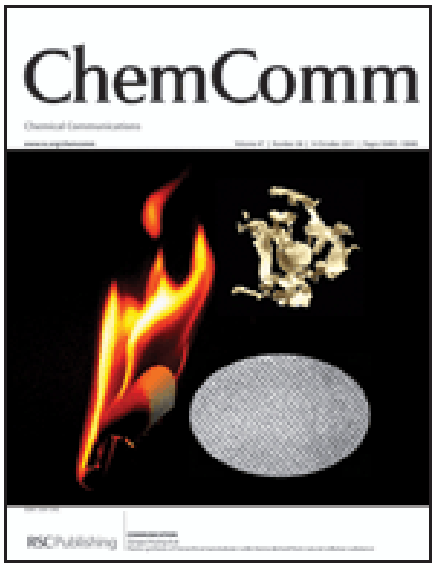
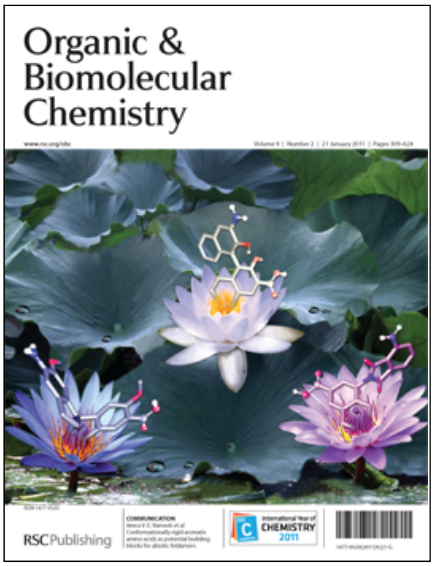




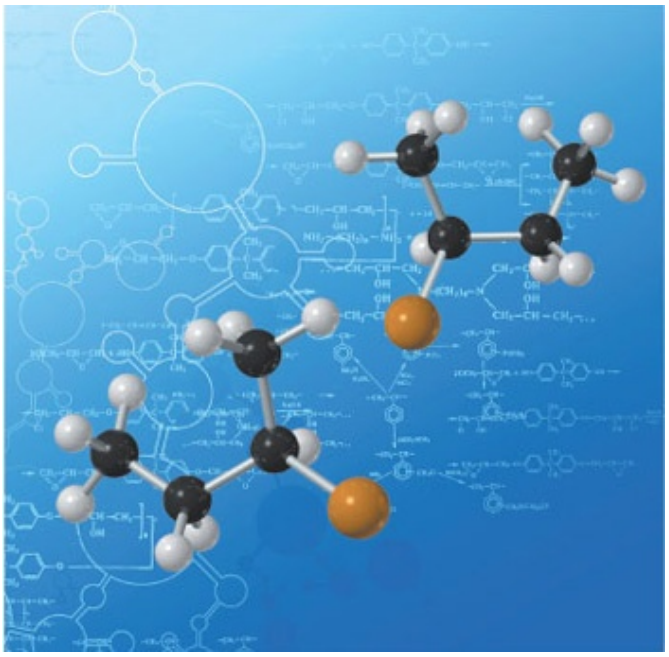
This article is part of the

Organocatalysis

web themed issue

Guest editors: Professors Keiji Maruoka, Hisashi Yamamoto, Liu-Zhu Gong and Benjamin List

All articles in this issue will be gathered together online at www.rsc.org/organocatalysis



Organic & Biomolecular Chemistry

Dynamic Article Links

Cite this: *Org. Biomol. Chem.*, 2012, **10**, 1565

www.rsc.org/obc

PAPER

Asymmetric organocatalytic formation of protected and unprotected tetroses under potentially prebiotic conditions†‡

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

Received 26th October 2011, Accepted 10th November 2011
DOI: 10.1039/c1ob06798b

Esters of proteinogenic amino acids efficiently catalyse the formation of erythrose and threose under potentially prebiotic conditions in the highest yields and enantioselectivities yet reported. Remarkably while esters of (L)-proline yield (L)-tetroses, esters of (L)-leucine, (L)-alanine and (L)-valine generate (D)-tetroses, offering the potential to account for the link between natural (L)-amino acids and natural (D)-sugars. The effect of pH and NaCl on the yields and enantioselectivities was also investigated and was shown to be significant, with the optimal enantioselectivities occurring at pH 7.

Introduction

One of the fundamental questions in the chemical and biological sciences is how did the reasonably complex building blocks of life, such as carbohydrates, form in the absence of any biological processes; and how did one enantiomer of these molecules come to dominate? In an important recent study, Sutherland has shown that RNA nucleotides can be prepared under plausible prebiotic conditions,[1] but this work requires glyceraldehyde to be present as a potential starting material. It has been suggested that the autocatalytic formation of glyceraldehyde from glycolaldehyde and formaldehyde in the presence of $Ca(OH)_2$ or other group II hydroxides (the formose reaction)[2] could be a possible route, as higher carbohydrates such as the tetroses and the pentoses have been detected in the reaction.[3,4] However, the formose reaction is inefficient in the formation of these higher carbohydrates as the basic reaction conditions lead to decomposition. Furthermore, it does not explain the emergence of homochirality in higher carbohydrates, although the recent work of Blackmond and Breslow provides insight into this problem.[5]

An alternative mechanism for the formation of carbohydrates, and one which can in theory generate enantiomeric enriched products, is the organocatalytic dimerization and trimerization of glycolaldehyde in water. However, this process is very inefficient in terms of both the yield and enantioselectivity of the carbohydrates formed. In a report by Pizzarello and Weber they demonstrated that the non-proteinogenic amino acid isovaline (100% ee) was able to catalyze the dimerization of glycolaldehyde in water, generating threose and erythrose products with enantiomeric excesses between 5–12%, although no yields were given.[6a] They also reported that dipeptides were also capable of catalyzing this reaction over extended reaction times at pH = 5.4 in higher enantioselectivities and yields.[6b] Zinc-proline complexes have also been reported to catalyze this dimerization and trimerization reaction in water, although the yields were low and no enantioselectivities were reported.[7]

Recently the pioneering studies of Barbas, List, MacMillan and Cordova have all independently shown that high enantioselectivities can be achieved in the aldol dimerization of aldehydes, including protected glycolaldehyde, in organic solvents using proline (100% ee) as a catalyst.[8–14] However, these reactions are nowhere near as successful in water, which must be viewed as the solvent for any prebiotic formation of carbohydrates.[12,13] In addition to this work, several groups have reported studies in which water was added to an organocatalytic aldol reaction in organic solvents.[15–26] In most cases that we are aware of the amount of water introduced is a relatively small quantity when compared to either the organic solvent or one of the aldol partners. Nonetheless, these studies have sparked some interesting debates on the nature of 'aqueous reactions'.[22,23] In a departure from the use of amino acids, Janda has shown that nicotine[24] and C2-aromatic substituted pyrrolidines are effective aldol-catalysts in water. Interestingly they showed that the best catalysts have electron-withdrawing substituents on the aromatic portion, although so far only racemic studies have been reported.[25]

We were intrigued by these reports and the lack of any efficient asymmetric organocatalytic process for the formation of

[a]*Department of Chemistry, University of York, Heslington, York, North Yorks, UK, YO10 5DD. E-mail: paul.clarke@york.ac.uk*
[b]*School of Chemistry, University of Nottingham, University Park, Nottingham, Notts, UK, NG7 2RD. E-mail: chris.hayes@nottingham.ac.uk*

†Electronic supplementary information (ESI) available: Experimental procedures for the preparation of compounds not detailed in the experimental section. Spectroscopic characterization of compounds not detailed in the experimental section. See DOI: 10.1039/c1ob06798b

‡This article is part of the joint ChemComm–Organic & Biomolecular Chemistry 'Organocatalysis' web themed issue.

carbohydrates in water, and so we set out to develop a simple, organocatalytic, high yielding, enantioselective and potentially prebiotic formation of tetroses in water. In light of the fact that proline is an excellent catalyst for the desired aldol dimerization in organic solvents, we wondered if simple modifications could be made to proline to enable it to efficiently catalyze the reaction in water. It is known that the carboxylic acid group of proline plays a key hydrogen-bonding role during the catalytic aldol reaction in organic solvents,[8] and this interaction will almost certainly be lost in aqueous conditions, which no doubt accounts for its lack of catalytic activity in water. We wondered whether simple esters of proline could act as aldol catalysts in water, as the resulting pyrrolidine would bear an electron-withdrawing substituent at C2, similar to the aromatic substituted catalysts of Janda.

Results and discussion

Our initial studies[27] focused on the dimerization of TIPS-protected glycolaldehyde. We chose this as our initial substrate for several reasons. Firstly, this aldehyde had been used by MacMillan in his work on the organocatalytic formation of sugars in organic solvents, and as such the products had been fully characterized, including the analysis of the enantioselectivities.[13] Secondarily, the reactions of protected versions of glycolaldehyde would be easier to monitor and their products easier to isolate. Thirdly, the experience gained in studying the TIPS-protected glycolaldehyde **1** reaction would be invaluable for when we started to investigate the aldol reaction of unprotected glycolaldehyde. The initial catalysts we alighted upon were the commercially available methyl **3**, ethyl **4** and benzyl **5** esters of (L)-proline. Additionally, we also synthesized the heneicosyl **6** and icosyl **7** esters of (L)-proline as Barbas had demonstrated that increasing the local hydrophobic environment of the catalyst led to enhanced enantioselectivities in some of the aldol reactions he had carried out in mixed organic/aqueous systems.[16] The heneicosyl **6** and icosyl **7** esters were prepared by the EDCI mediated coupling of Boc-(L)-proline with the appropriate alcohol. With all five catalysts in hand we examined their efficiency at catalyzing the aqueous aldol dimerization of TIPS-glycolaldehyde **1** in water. We were pleased to find that these (L)-proline esters catalyzed the reaction efficiently, giving tetrose products in good isolated yields with moderate diastereoselectivity and enantioselectivity (Fig. 1, Table 1).

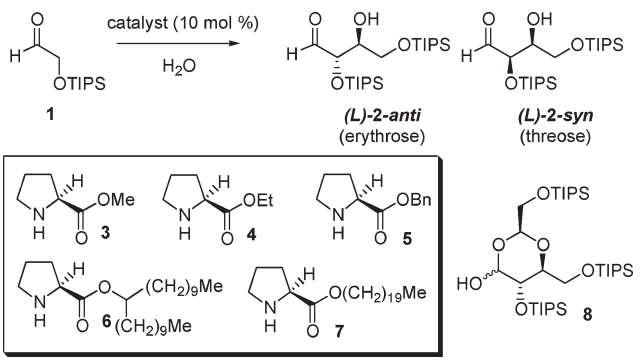


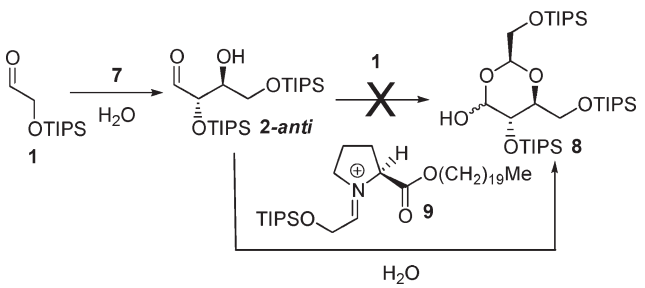
Fig. 1 (L)-proline ester catalyzed dimerization and trimerization of TIPS-glycolaldehyde in water.

Table 1 (L)-proline ester catalyzed formation of TIPS-protected tetroses in water

Entry	Catalyst	Major product[c]	Combined yield (%)	Ratio (*anti* : *syn*)[d]	% ee (*anti*)[e]
1[a]	**3**	(L)-**2**-*anti*	68	4.5 : 1	15
2[a]	**4**	(L)-**2**-*anti*	57	4 : 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**	(L)-**2**-*anti*	49	2 : 1	10
6[b]	**7**	**8**	40[f]	N/A[g]	17[h]

[a] Reaction time 48 h. [b] Reaction time 5 h. [c] Major enantiomer determined by correlation to the work of Northrup and MacMillan (ref. 13). [d] Determined by integration of the aldehyde resonances in the ^{1}H NMR. [e] See supporting information.† [f] Isolated yield of acetal **8**. [g] 10% of **2**-*anti* was also formed. [h] % ee of acetal **8**.

Initially the reactions were run for 48 h (entries 1–3), but that could be shortened to 5 h (entries 4, 5) without detriment to the yield or stereoselectivity of the reaction. It is worth noting that this level of enantioselectivity is slightly better than that reported by Pizzarello and Webber in their earlier work.[6a] Interestingly when catalyst **7** was employed (entry 6) the major product was not the tetrose product (the *anti*-aldol dimer accounted for only 10% of the product), but acetal **8**. Acetal **8** is effectively a trimer of TIPS-glycolaldehyde **1**, and its formation was unexpected. It is obviously produced by the reaction of the initially formed *anti*-aldol dimer **2**-*anti* and another equivalent of **1** by an acetalization process. While these acetals have been isolated before in organocatalytic aldol reaction in organic solvents, notably by MacMillan,[28] it is surprising that they form so readily and in such high yields in an aqueous media. As we do not see formation of acetal **8** with any of the other catalysts, we believe that **8** is formed from **2**-*anti* by reaction with the catalyst-derived iminium ion **9**, rather than by reaction with **1** directly (Scheme 1).



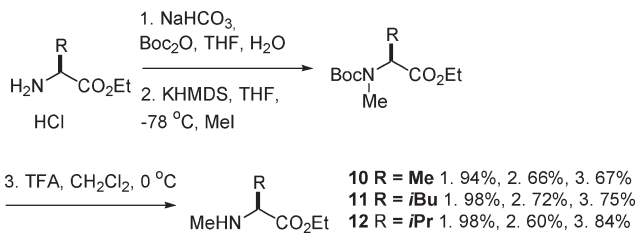
Scheme 1 (L)-Proline icosyl ester catalyzed formation of cyclic acetal **8**.

Having demonstrated that esters of (L)-proline were able to catalyze the aldol dimerization of **1**, we wondered if esters of other proteinogenic amino acids could also catalyze this reaction. As the key intermediate in the aldol dimerization reaction is the formation of a reactive enamine intermediate, we were concerned that other amino acid esters may not be as efficient as proline in this reaction as they contain primary amine groups which may stall the reaction due to the catalyst being trapped as an imine or hemiaminal. In order to obviate this possibility we decided to synthesize *N*-methyl derivatives of amino acid esters of (L)-alanine **10**, (L)-leucine **11** and (L)-valine **12** (Scheme 2).

Table 2 (L)-*N*-Methyl-amino acid ester catalyzed formation of TIPS-protected tetroses in water

Entry	Catalyst	Major product[a]	Combined yield (%)	Ratio (*anti* : *syn*)[b]	% ee (*anti*)[c]
1	**10**	(D)-**2**-*anti*	70	3 : 1	7
2	**11**	(D)-**2**-*anti*	80	1.5 : 1	17
3	**12**	(D)-**2**-*anti*	33	1 : 1	31

[a] Major enantiomer determined by correlation to the work of Northrup and MacMillan (ref. 13). [b] Determined by integration of the aldehyde resonances in the ^{1}H NMR. [c] See supporting information.†



Scheme 2 Synthesis of catalysts **10**, **11** and **12**.

Catalysts **10**, **11** and **12** were prepared by Boc-protection of the ethyl ester of the appropriate amino acid with Boc_2O followed by deprotonation of the carbamate nitrogen and trapping the anion with methyl iodide. The catalyst was then revealed by TFA-mediate deprotection which provided **10**, **11** and **12** in 42%, 53% and 49% overall yields respectively. These catalysts were then tested in the dimerization reaction of **1** in water over 5 h (Table 2).

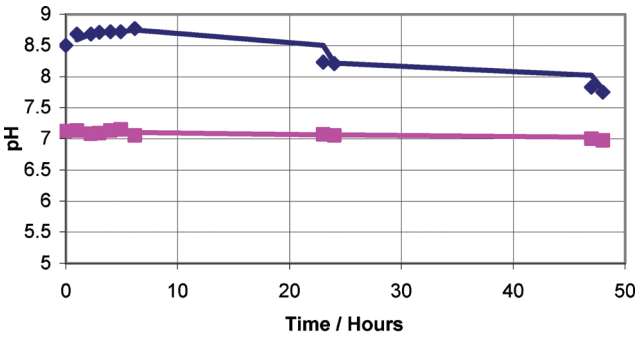
Catalysts **10** and **11** catalyzed the reaction in moderate to good yields and diastereoselectivities, with enantioselectivities of the same numerical magnitude as the (L)-proline ester catalysts **3–6** (entries 1 and 2). The (L)-valine derived catalyst **12** catalyzed the reaction in lower yield and diastereoselectivity, but with an enantioselectivity nearly double that of the (L)-leucine derived catalyst **11** (entry 3) or over four times that of the (L)-proline derived catalysts (Table 1). This is the largest % ee to date reported for the dimerization of **1** in water with an amino acid derivative as catalyst. Surprisingly however, it was the (D)-enantiomer of erythrose (D)-**2**-*anti* which was generated as the major enantiomer in these reactions. This is in contrast to the reaction run with catalysts derived from (L)-proline which generated (L)-erythrose (L)-**2**-*anti* as the major product. This switch from (L)-erythrose with (L)-proline catalysts to (D)-erythrose with (L)-alanine, (L)-leucine and (L)-valine marks an unexpected and significant switch in the enantiomer of the product generated.

Janda has pointed out that general base catalysis of the aldol dimerization may operate when amine-based catalysts are used in aqueous media,[22] and this could well lead to a reduction in the enantioselectivity seen. We wished to explore the influence of pH on the reaction and decided to monitor the pH of the aldol dimerization reaction and to run the reaction at both pH = 6, which should be optimal for enamine formation, and pH = 7. At pH = 6 there should be no general base catalysis, however, some acid catalysis may operate. At pH = 7 there should be no general base or acid catalysis and so the yields and enantioselectivities should result from an enamine catalyzed process. The benzyl ester of (L)-proline **5** was chosen as the catalyst for these studies.

Table 3 (L)-Amino acid ester catalyzed formation of TIPS-protected tetroses in water and pH = 7 buffer

Entry	Catalyst	Major product[b]	Combined yield (%)	Ratio (*anti* : *syn*)[c]	% ee (*anti*)[d]
1[a]	**5**	(L)-**2**-*anti*	70	1.5 : 1	47
2[a]	**6**	(L)-**2**-*anti*	52	5.5 : 1	46
3[a]	**7**	**8**	32[e]	N/A	23[f]
4[a]	**11**	(D)-**2**-*anti*	79	1.5 : 1	57
5[a]	**12**	(D)-**2**-*anti*	40	1.5 : 1	79

[a] Reaction time 5 h. [b] Major enantiomer determined by correlation to the work of Northrup and MacMillan (ref. 13). [c] Determined by integration of the aldehyde resonances in the ^{1}H NMR. [d] See supporting information.† [e] Isolated yield for acetal **8**. [f] % ee of acetal **8**.



Blue line = unbuffered H_2O; Pink line = pH 7 buffer.

Fig. 2 pH of aldol reaction of TIPS-glycolaldehyde **1** catalyzed by (L)-proline benzyl ester.

The first reaction repeated the conditions reported in Table 1, entry 5 and the pH was monitored over the 48 h period. From the pH profile (Fig. 2) it can be clearly seen that the pH at the start of the reaction is basic (pH = 8.5) and that it increases to a maximum (pH = 8.8) after approximately 5 h. After this time the pH drops steadily to a final pH = 7.8 after 48 h. These observations clearly demonstrated that a general base mediated aldol reaction could be in operation under these conditions. We next decided to run the dimerization of **1** in the presence of KH_2PO_4/NaOH buffered to pH = 7 and use each of the available catalysts **3–7** and **10–12** (Table 3).

We were please to find that under these buffered conditions we still obtained good yields of tetrose products **2**-*syn* and **2**-*anti*. More importantly, however, was the increase in enantioselectivity of the reactions, with the % ee of each of the buffered reactions being substantially higher than the comparative unbuffered reactions. In the case of catalysts **5** and **6** the % ee increased from 18% to 47% to 10% to 46% respectively (entries 1 and 2). The biggest increase in enantioselectivity was seen with catalyst **11** which in the unbuffered reaction gave the product with a 17% ee and in the pH = 7 buffered reaction generated product with a 57% ee (entry 4). The reaction with the highest enantioselectivity used (L)-valine derived catalyst **12** which in the pH = 7 buffered reaction returned a product with an impressive 79% ee (entry 5), compared to 31% ee from the unbuffered reaction. Significantly, (L)-proline catalysts **5** and **6** still produced (L)-erythrose as the major enantiomer, while (L)-leucine and (L)-valine catalysts still produced (D)-erythrose as the

Table 4 (L)-Amino acid ester catalyzed formation of TIPS-protected tetroses in water and pH = 6 buffer

Entry	Catalyst	Major product[b]	Combined yield (%)	Ratio (:*syn*)[c]	% ee (*anti*)[d]
1[a]	**5**	(L)-**2**-*anti*	33	5 : 1	22
2[a]	**6**	(L)-**2**-*anti*	30	8 : 1	20
3[a]	**7**	(L)-**2**-*anti*	24	1 : 1	30
4[a]	**11**	(D)-**2**-*anti*	33	1 : 1	16

[a] Reaction time 5 h. [b] Major enantiomer determined by correlation to the work of Northrup and MacMillan (ref. 13). [c] Determined by integration of the aldehyde resonances in the ^{1}H NMR. [d] See supporting information.†

major enantiomer, thus showing that while the level of the asymmetric induction was influenced by pH, the absolute sense of the reaction was not. It is worth noting that the use of catalyst **7** under pH = 7 buffered conditions still generated acetal **8** as the major product (entry 3) in a similar yield and slightly higher % ee to that seen in the unbuffered reaction. These results indicate that a general base-mediated condensation may well be occurring in the unbuffered reaction and could account for the reduced enantioselectivities. They also suggest that the enamine-mediated reaction is at least as efficient (in terms of yield) as the general base-mediated reaction. It is worth noting too, that these results represent the highest enantioselectivity (79% ee) seen in the single amino acid residue catalyzed dimerization of a protected glycolaldehyde derivative in an aqueous environment.

We next decided to see what effect running the reactions at pH = 6 would have on the yield, diastereoselectivity and enantioselectivity of tetrose **2**-*syn*/**2**-*anti* formation. The reactions were buffered to pH = 6 with KH_2PO_4/NaOH and run for 5 h (Table 4) as in the previous trials in unbuffered solution and when buffered at pH = 7. As can be seen from Table 4, the yields of all of the reactions were substantially lower at pH = 6 than at the other pHs investigated. This came as a surprise as we envisaged a yield increase due to pH = 6 being the accepted optimal pH for enamine formation. However it was noted that there was a significant amount of unidentified silyl residue in the ^{1}H NMR of the crude reactions, so we assume that degradation of either the starting material **1** or products **2**-*syn*/**2**-*anti* was occurring over the reaction time. In the cases of catalysts **5** and **6** (entries 1 and 2) there was a significant increase in the diastereoselectivity of the reaction with more of the erythrose product **2**-*anti* produced. The enantioselectivities were lower than those that had been seen for the pH = 7 buffered reaction, but were in general higher than for the unbuffered reaction. No acetal **8** formation was seen at pH = 6 when catalyst **7** was used, with the reaction returning tetrose products in yields and enantioselectivities comparable to the other catalysts in this set of reactions.

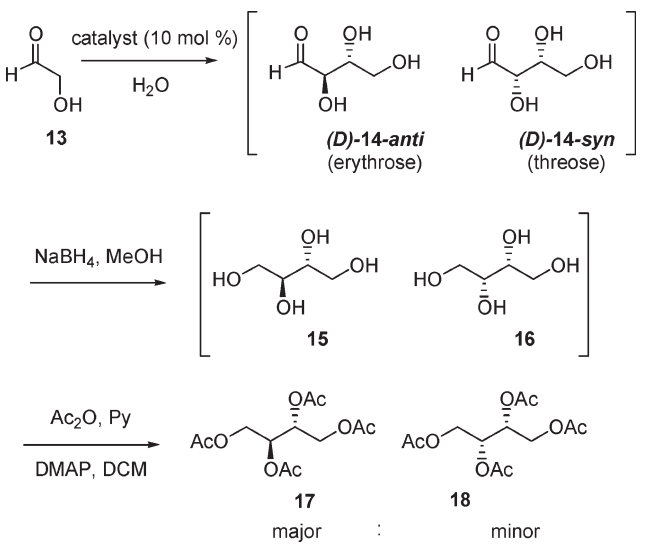
Table 5 (L)-Amino acid ester catalyzed formation of TIPS-protected tetroses in unbuffered brine

Entry	Catalyst	Major product[b]	Combined yield (%)	Ratio (*anti* : *syn*)[c]	% ee (*anti*)[d]
1[a]	**5**	(L)-**2**-*anti*	8	1 : 1	—[e]
2[a]	**6**	**8**[f]	41[g]	N/A[f]	15[h]
3[a]	**7**	(L)-**2**-*anti*	57	1 : 1	23
4[a]	**10**	(D)-**2**-*anti*	31	1 : 1	23
5[a]	**11**	(D)-**2**-*anti*	26	1 : 1	16

[a] Reaction time 5 h. [b] Major enantiomer determined by correlation to the work of Northrup and MacMillan (ref. 13). [c] Determined by integration of the aldehyde resonances in the ^{1}H NMR. [d] See supporting information.† [e] Not determined. [f] The reaction also produced **2**-*anti*/**2**-*syn* in a 1.0 : 1.0 *anti* : *syn* ratio where **2**-*anti* could be isolated in 18% yield, 0% ee (*anti*). [g] Isolated yield for acetal **8**. [h] % ee of acetal **8**.

The final set of reaction conditions we examined was running the aldol dimerization of **1** in unbuffered brine in an attempt to mimic sea water[29] as the reaction medium (Table 5). In the case of catalyst 5 (entry 1) the yield was too low to enable us to determine the enantioselectivity of the reaction, however, the ratio of erythrose **2**-*anti*/threose **2**-*syn* was 1 : 1. The trend of (L)-proline derived catalysts giving (L)-erythrose and the major product and (L)-alanine and (L)-leucine catalysts giving (D)-erythrose as the major enantiomer continued (entries 4 and 5). Thus demonstrating that for all the conditions investigated there is a link between acyclic (L)-amino acids and (D)-tetroses. This is significant and consistent with the work on the organocatalytic formose reaction recently reported by Breslow.[5c] Interestingly under these conditions it is catalyst **6**, not **7**, which generates acetal **8** as the major product in 41% yield and 15% ee. A 1 : 1 mixture erythrose **2**-*anti* and threose **2**-*syn* was also isolated from this reaction, although the erythrose **2**-*anti* was racemic. Considering that the acetal **8** had an ee of 15% this provides further evidence of the intermediacy of an iminium (such as **9**), as a kinetic resolution of the erythrose **2**-*anti* seems to be occurring.

With studies on the dimerization of **1** and hence the synthesis of TIPS-protected erythrose **2**-*anti* and threose **2**-*syn* complete we decided to investigate whether our catalysts could promote the dimerization of unprotected glycolaldehyde **14**, and hence lead directly to the formation of erythrose **14**-*anti* and threose **14**-*syn* (Scheme 3). We chose to use catalysts **11** and **12** as they had consistently provided the highest enantioselectivities and should generate the (D)-enantiomers of the product tetroses. We also decided to investigate the tetrose-forming reaction under each of the conditions we had studied for the formation of the TIPS-protect tetroses. Hence the reactions were run with catalysts **11** and **12** for 5 h in unbuffered aqueous media, water



Scheme 3 Organocatalytic synthesis of erythrose and threose under aqueous conditions.

Table 6 (L)-amino acid ester catalyzed formation of unprotected (D)-tetroses

Entry	Medium	Catalyst	Yield of **17** (%)[b]	Yield of **18** (%)[b]	Product enantiomer[c]	% ee **18**[d]
1[a]	unbuffered	**11**	2.4	1.6	(D)-**18**	25
2[a]	pH 7	**11**	7.1	0.9	(D)-**18**	66
3[a]	pH 6	**11**	2.5	0.5	(D)-**18**	43
4[a]	brine	**11**	2.0	2.0	(D)-**18**	20
5[a]	unbuffered	**12**	0.0	0.0	N/A	N/A
6[a]	pH 7	**12**	3.0	2.0	(D)-**18**	40
7[a]	pH 6	**12**	3.0	3.0	(D)-**18**	14
8[a]	brine	**12**	9.8	2.2	(D)-**18**	30

[a] Reaction time 5 h. [b] After reduction and acylation. [c] Major enantiomer determined by correlation to the work of Cordova (ref. 17). [d] Determined by GC.

buffered to pH = 7, water buffered to pH = 6 and unbuffered brine (Table 6). However, due to the difficulty inherent in directly assaying the yield, diastereo- and enantioselectivity of the tetrose products we adopted the analysis protocol reported by Cordova.[17] This involved reduction of the tetrose products with $NaBH_4$ in MeOH followed by acylation with acetic anhydride in pyridine/CH_2Cl_2 with DMAP and analysis by chiral GC. The unfortunate consequence of this procedure is that upon reduction the erythrose **14**-*anti* is converted to erythritol **15** which is a *meso*-compound, and so only the enantiomeric excess of the threitol-derived tetraacetate **18**, and hence of threose **14**-*syn* could be determined.[30]

As can be seen from Table 6, in all cases the major enantiomer of the product formed was the (D)-enantiomer, continuing the trend seen in our earlier work on the protected glycolaldehyde. The (L)-leucine derived catalyst **11** generally gave products with higher % ee than the (L)-valine derived catalyst **12** (entries 1–4 to entries 5–8). Dimerization reactions run at pH = 7 gave the highest % ee out of all the conditions investigated (entries 2 and 6), with the highest overall % ee of 66% arising from the (L)-leucine derived catalyst **11** at pH = 7 (entry 2). To date this is the highest % ee for the dimerization of glycolaldehyde by a catalyst containing a single amino acid under aqueous conditions. As in the case of the dimerization of the protected glycolaldehyde **1**, those reactions run at pH = 6 and in brine had lower % ee values (entries 3, 4, 7 and 8). It is possible that this is due to competing general base or acid catalysis. Surprisingly, when the dimerization reaction was run unbuffered with catalyst **12**, no product could be isolated (entry 5). We attribute the lower yields of these unprotected reactions to the difficulty in product manipulation and isolation. It is probable that not all of the tetrose products were reduced and converted to the tetraacetates thus resulting in loss of material to the aqueous phase. Although it is worth noting that the (L)-valine derived catalyst **12** often gives lower isolated yields of products than the (L)-leucine derived catalyst **11**.

Conclusions

We have developed a simple system for the formation of threose and erythrose, catalyzed by esters of proteinogenic amino acids, under aqueous and potentially prebiotic conditions. This has resulted in the highest enantioselectivities reported to date for the formation of these simple tetroses. Studies on the effect of pH on the reaction indicate that the tetroses are most likely formed *via* an enamine-mediated process but that in systems where the pH <7 or pH >7 the enantioselectivity of the dimerization could be eroded by general acid or general base catalysis respectively. Fascinatingly and significantly while esters of (L)-proline generated tetrose products in the unnatural (L)-enantiomeric series, esters of acyclic (L)-amino acids (valine, alanine and leucine) all generated tetrose products in the natural (D)-enantiomeric series. This offers one potential explanation to account for the relationship between (L)-amino acids and (D)-sugars in nature. We have also noted that while hexoses, the products of aldol trimerization, are not formed in these reactions, 'hexose-like' acetal trimers are formed efficiently with some (L)-proline derived catalysts.

Experimental section

General procedure for the amino acid ester catalyzed dimerization of 1

Amino acid ester (0.129 mmol) was added to 2-(triisopropylsilyloxy)acetaldehyde **1** (280 mg, 1.29 mmol) in either water, or pH 7 phosphate buffer or pH 6 phosphate buffer or brine (5 mL). After 5 h the reaction mixture was extracted with chloroform (3 × 10 mL). The organic extracts were combined, dried with magnesium sulfate, filtered and concentrated *in vacuo* to provide the crude reaction mixture. This was then purified by flash column chromatography using Merck TLC grade silica gel with silica/alumina binder supplied by Aldrich (15 : 1 pentane : diethyl ether) to provide clean **2**-*anti* (in quantities suitable for the analysis of enantioselectivity) and a mixture of **2**-*syn*/**2**-*anti* as a clear, colourless oil (See Tables 1–5 for combined yields of **2**-*anti* and **2**-*syn*/**2**-*anti* product mix and diastereomeric ratios of the crude reaction mixture). The data was found to be in accordance with the literature.[12] ^{1}H NMR (400 MHz, $CDCl_3$): δ (*anti*): 1.03–1.08 (m, 42H, $SiCH(CH_3)_2$); 2.39 (d, 5.5 Hz, 1H, OH); 3.75–3.85 (m, 2H, CH_2); 3.95–3.99 (m, 1H, $CHOHCH_2$); 4.24 (dd, 4.0 Hz, 2.0 Hz, 1H, CHOHCHO); 9.68 (d, 2.0 Hz, 1H, CHO); (*syn*): 1.03–1.08 (m, 42H, $SiCH(CH_3)_2$); 2.74 (d, 10.0 Hz, 1H, OH); 3.76–3.92 (m, 2H, CH_2); 3.95–3.98 (m, 1H, $CHOHCH_2$); 4.28 (dd, 5.0 Hz, 2.0 Hz, 1H, CHOHCHO); 9.74 (d, 1.5 Hz, 1H, CHO);^{13}C NMR (100 MHz, $CDCl_3$): δ (*anti*): 11.9, 12.4, 18.0, 62.8, 74.4, 79.0, 202.4; (*syn*): 11.9, 12.4, 18.0, 62.2, 74.4, 78.1, 203.8.

General procedure for conversion of 2-*syn*/2-*anti* to the *para*-nitrobenzoate and the determination of its enantiomeric excess

A sample of (*anti*)-3-hydroxy-2,4-bis(triisopropylsilyloxy) butanal **2**-*anti* (10.7 mg, 0.0247 mmol) was added to a solution of 4-nitrobenzoyl chloride (11.3 mg, 0.0609 mmol) and 4-(dimethylamino)pyridine (0.664 mg, 0.00543 mmol) in dichloromethane (0.5 mL) at 0 °C under argon. Triethylamine (0.0172 mL, 0.124 mmol) was then added to the reaction mixture. After 3 h methanol (0.5 mL) was added to the solution, followed by sodium borohydride (9.34 mg, 0.247 mmol). After 1 h the reaction mixture was allowed to warm to room temperature. After a further 2 h the solution was diluted with dichloromethane (5 mL) and washed with saturated sodium bicarbonate

solution (5 mL). The aqueous layer was extracted with dichloromethane (3 × 2 mL). The organic extracts were then combined, dried with magnesium sulfate, filtered and concentrated *in vacuo* to give the crude acylated and reduced product. Purification by preparative thin-layer chromatography (4 : 1 pentane : diethyl ether) provided the 1-hydroxy-3-*p*-nitrobenzoate derivative (2.8 mg, 0.0048 mmol, 19%) as a colourless oil. The enantiomeric excess was then determined by chiral shift NMR analysis using europium tris[3-(trifluoromethylhydroxymethylene)-(+)-camphorate] (0.9 mg, 0.0010 mmol) in deuterated chloroform (0.7 mL). Only HPLC data provided in the literature,[12] full data provided below; IR (film) 2945, 2891, 2867, 1724, 1607, 1531, 1462, 1384, 1349, 1320, 1278, 1103, 1059, 1015, cm^{-1}; ^1H NMR (270 MHz, $CDCl_3$): δ 0.95 (m, 42H, $SiCH(CH_3)_2$); 3.55–3.81 (m, 2H, CH_2OH); 3.90–4.09 (m, 2H, $CH_2OSiCH(CH_3)_2$); 4.21 (q, 4.0 Hz, 1H, $CHCH_2OH$); 5.32 (q, 4.0 Hz, 1H, CHOOCAr); 8.19 (dd, 9.0 Hz, 13.0 Hz, 4H, $C_6H_4NO_2$); ^{13}C NMR (100 MHz, $CDCl_3$): δ 11.8, 12.5, 17.9, 18.0, 18.1, 27.0, 30.3, 61.7, 63.4, 72.6, 76.7, 123.5, 130.8,135.6, 150.6, 164.3; HRMS (ESI) exact mass calcd for $[M + H]^+$ ($C_{29}H_{54}NO_7Si_2$) requires *m*/*z* 584.3433, found *m*/*z* 584.3436.

General procedure for the amino acid ester catalyzed dimerization of glycolaldehyde 13

Glycolaldehyde dimer **13** (240 mg, 2.00 mmol) was added to a stirred mixture of amino acid ester (0.10 mmol) in either water or pH 7 phosphate buffer, or pH 6 phosphate buffer or brine (3 mL). After 5 h the reaction mixture was concentrated *in vacuo*, and re-dissolved in methanol (3 mL) at 0 °C. Sodium borohydride (152 mg, 4.00 mmol) was then added carefully and the reaction kept at 0 °C for 3 h, after which it was allowed to warm to room temperature. After a further 15 h, the reaction was cooled to 0 °C and quenched with 2 M hydrochloric acid (3 mL). It was then concentrated *in vacuo* and re-dissolved in dichloromethane (5 mL). Pyridine (1 mL) was then added, followed by 4-dimethylaminopyridine (1.3 mg, 0.01 mmol) then acetic anhydride (3 mL). The reaction mixture was stirred for 7 h, then washed with water (10 mL) and extracted with dichloromethane (3 × 5 mL). The separated and combined organics were then washed with 1 mol dm^{-3} hydrochloric acid (10 mL), brine (10 mL) and finally water (10 mL). The organic layer was then dried with magnesium sulfate and concentrated *in vacuo* to give the crude reduced and acylated derivative tetroses (see Table 6). These were purified by flash column chromatography using silica gel 60 (220–240 mesh) (8 : 2 hexane : ethyl acetate) to give the separated acylated tetrols **17** and **18** as a colourless oils (individual yields provided in Table 6). The chiral product **18** was then dissolved in ethyl acetate and analysed by chiral-phase GC analysis using the conditions provided in the literature.[17,31] GC tetra-acetylated threitol: (CP-Chirasil-Dex CB); T_{inj} = 250 °C, T_{det} = 275 °C, flow = 1.5 mL min^{-1}, t_i = 100 °C (10 min), (100 °C min^{-1}) t_f = 200 °C (40 min): (L)-isomer: t_R = 36.33 min; (D)-isomer: t_R = 36.79 min.

Acknowledgements

We thank the University of York/EPSRC DTA (LB), the University of Nottingham (MEV) and the ERASMUS exchange scheme (MZ) for funding, Prof. Ian Fairlamb for use of the GC.

References

1 M. W. Powner, B. Gerland and J. D. Sutherland, *Nature*, 2009, **459**, 239.
2 (*a*) A Butlerow, *Justus Liebigs Ann. Chem.*, 1861, **120**, 295; (*b*) R. Breslow, *Tetrahedron Lett.*, 1959, **1** (21), 22.
3 A. Ricardo, M. A. Carrigan, A. N. Olcott and S. A. Benner, *Science*, 2004, **303**, 196.
4 J. B. Lambert, S. A. Gurusamy-Thangavelu and K. Ma, *Science*, 2010, **327**, 984.
5 (*a*) J. E. Hein, E. Tse and D. G. Blackmond, *Nat. Chem.*, 2011, **3**, 704; (*b*) R. Breslow, *Tetrahedron Lett.*, 2011, **52**, 2028; (*c*) R. Breslow and Z.-L. Chen, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 5723.
6 (*a*) S. Pizzarello and A. L. Weber, *Science*, 2004, **303**, 1151; (*b*) A. L. Weber and S. Pizzarello, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 12713.
7 J. Kofoed, J.-L. Reymond and T. Darbre, *Org. Biomol. Chem.*, 2005, **3**, 1850.
8 B. List, R. A. Lerner and C. F. Barbas III, *J. Am. Chem. Soc.*, 2000, **122**, 2395.
9 K. Sakthivel, W. Notz, T. Bui and C. F. Barbas III, *J. Am. Chem. Soc.*, 2001, **123**, 5260.
10 B. List, L. Hoang and H. J. Martin, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 5839.
11 B. List, *Tetrahedron*, 2002, **58**, 5573.
12 A. B. Northrup, I. K. Mangion, F. Hettche and D. W. C. MacMillan, *Angew. Chem., Int. Ed.*, 2004, **43**, 2152.
13 A. B. Northrup and D. W. C. MacMillan, *Science*, 2004, **305**, 1752.
14 A. Córdova, M. Engqvist, I. Ibrahem, J. Casas and H. Sundén, *Chem. Commun.*, 2005, 2047.
15 A. Córdova, W. Notz and C. F. Barbas III, *Chem. Commun.*, 2002, 3024.
16 N. Mase, Y. Nakai, N. Ohara, H. Yoda, K. Takabe, F. Tanaka and C. F. Barbas III, *J. Am. Chem. Soc.*, 2006, **128**, 734.
17 A. Córdova, I. Ibrahem, J. Casas, H. Sundén, M. Engqvist and E. Reyes, *Chem.–Eur. J.*, 2005, **11**, 4772.
18 J. Mlynarski and J. Paradowska, *Chem. Soc. Rev.*, 2008, **37**, 1502.
19 F. Rodríguez-Llansola, J. F. Miravet and B. Escuder, *Chem. Commun.*, 2009, 7303.
20 M. De Nisco, S. Pedatella, H. Ullah, J. H. Zaidi, D. Naviglio, Ö. Özdamar and R. Caputo, *J. Org. Chem.*, 2009, **74**, 9562.
21 N. Mase, N. Noshiro, A. Mokuya and K. Takabe, *Adv. Synth. Catal.*, 2009, **351**, 2791.
22 A. P. Brogan, T. J. Dickerson and K. D. Janda, *Angew. Chem., Int. Ed.*, 2006, **45**, 8100.
23 Y. Hayashi, *Angew. Chem., Int. Ed.*, 2006, **45**, 8103.
24 T. J. Dickerson and K. D. Janda, *J. Am. Chem. Soc.*, 2002, **124**, 3220.
25 C. J. Rogers, T. J. Dickerson, A. P. Brogan and K. D. Janda, *J. Org. Chem.*, 2005, **70**, 3705.
26 (*a*) Y. Hayashi, T. Sumiya, J. Takahashi, H. Gotoh, T. Urushima and M. Shoji, *Angew. Chem., Int. Ed.*, 2006, **45**, 958; (*b*) S. Aratake, T. Itoh, T. Okano, T. Usui, M. Shoji and Y. Hayashi, *Chem. Commun.*, 2007, 2524; (*c*) S. Aratake, T. Itoh, T. Okano, N. Nagae, T. Sumiya, M. Shoji and Y. Hayashi, *Chem.–Eur. J.*, 2007, **13**, 10246; (*d*) Y. Hayashi, S. Aratake, T. Itoh, T. Okano, T. Sumiya and M. Shoji, *Chem. Commun.*, 2007, 957.
27 L. Burroughs, M. E. Vale, J. A. R. Gilks, H. Forintos, C. J. Hayes and P. A. Clarke, *Chem. Commun.*, 2010, **46**, 4776.
28 I. K. Mangion, A. B. Northrup and MacMillan, *Angew. Chem., Int. Ed.*, 2004, **43**, 6722.
29 The precise composition of sea water on the prebiotic (Hadean) Earth (4.3–3.8 Ga) is a matter of some debate as there is no geological record of rocks of that age. There is, however, a consensus view that the oceans did contain sodium chloride, much the same as modern oceans. See: J. W. Morse and F. T. Mackenzie, *Aquat. Geochem.*, 1998, **4**, 301.
30 The reduction–acetylation protocol was performed independently on samples of enantiomerically pure erythrose and threose to provide reference samples for chiral-GC analysis. In addition, these studies confirmed that threose and erythrose do not interconvert under the reaction conditions.
31 Spectroscopic data for the tetraacetate standards are presented in the supporting information†.

