RS•C

Chemistry NOW

This is a series of four leaflets which present modern aspects of chemistry in a way accessible to school students and directly usable by teachers. Each leaflet consists of four pages of information interspersed with questions to test student's understanding of what they are reading, to help them to link what they have read to the chemistry they already know and to help them to understand the text.

The leaflets could be used to support existing workschemes, to develop comprehension skills or as meaningful exercises to be used in the case of teacher absences (planned or unplanned).

The leaflets are:

•Chemistry and sport

This is aimed at 14–16 year olds and deals with the chemistry of aerobic and anaerboic respiration in the context of athletics and looks at a number of ways in which athletes can manipulate (legally!) the chemistry of this process to their advantage by monitoring the concentration of lactic acid in their blood.

•Chemistry of the atmosphere

This is aimed at 14–16 year olds. This looks at the way that the Earth's present atmosphere has evolved from possible earlier atmospheres. Some of the available evidence for different scenarios is presented and critically discussed.

•Computational chemistry

This is aimed at the post-16 age group. It presents a case study of the development of derivatives of cinnamic acid as a repellent to dissuade birds from eating crops treated with it. It explains how chemists develop relationships between structural features and particular types of activity and how computer modelling programmes are used in this work.

•Combinatorial chemistry

This is also aimed at the post-16 age group. Combinatorial chemistry is a group of techniques for synthesising large arrays of related chemicals. These can be easily automated by the use of robot syringes controlled by computers to carry out repetitive processes. The resulting arrays of chemicals called 'libraries' can then be screened for potential drug activities. Combinatorial chemistry is increasingly being used by pharmaceutical companies in their search for new drugs.

Electron distribution of cinnamamide

Answers

For the use of teachers, answers to the questions on the leaflets are presented overleaf.

RS•C

Combinatorial chemistry –

the future for drug discovery?

•Finding new drugs

Drug discovery has been compared to finding a needle in a haystack. The needle is an effective and safe drug and the haystack is the vast range of chemical compounds that have been, or could be, made.

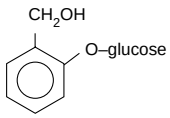
Of course, pharmaceutical chemists do not synthesise compounds at random. They usually start with a so-called lead compound (*ie* a compound which leads them onward, not a compound of the metal lead), and make and test compounds with structures closely related to it. This lead compound will have shown some activity as a drug and the chemists hope that they can find a derivative of it which is better – is more effective as a drug, has fewer side effects, is cheaper to make *etc*. A good example of the lead compound approach is the development of the drug aspirin (Box 1).

Compound libraries

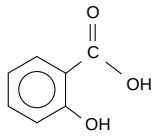
Collections of compounds to be screened for drug activity are called 'libraries'. A drug company might have a library consisting of maybe a million compounds which it has synthesised or bought in. Of these only a handful will show sufficient promise as possible drugs to be developed further. Using traditional methods of chemical synthesis the potential drug substances are synthesised one at a time, purified, and then, after their structures have been confirmed, are tested for their effectiveness as drugs. So it could take many years to build up such a library.

THE DEVELOPMENT OF ASPIRIN **Box 1**

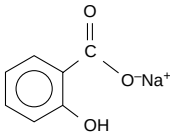
For over 2000 years, extracts of willow bark have been used to treat pain and fever. In the 1840s, the compound salicin was found to be the active ingredient. This was the lead compound.



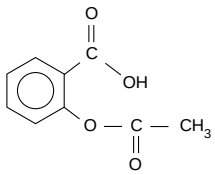
In the 1870s, the related compound salicylic acid (2-hydroxybenzenecarboxylic acid) was used in patients to good effect but it caused irritation and bleeding of the intestinal tract.



Later the sodium salt of this acid (sodium 2-hydroxybenzenecarboxylate) was used. This too was effective but caused vomiting and had a dreadful taste.



In the 1890s, 2-ethanoyloxybenzenecarboxylic acid – another derivative of salicin – was synthesised. This was just as effective but caused less stomach irritation and tasted less unpleasant. This is the drug we know as aspirin.



•The principles of combinatorial chemistry

Over the past few years several techniques called collectively 'combinatorial chemistry' have been developed which enable chemists to prepare libraries of thousands of related compounds quickly.

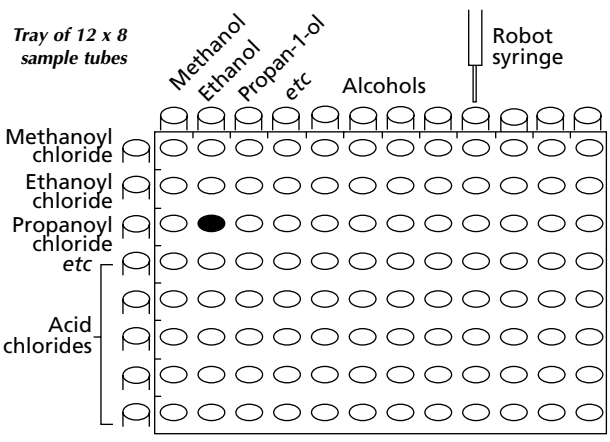
These techniques use automated methods called 'parallel synthesis' (Box 2).

PARALLEL SYNTHESIS **Box 2**

The principle of parallel synthesis is shown in the following example. Alcohols react with acid chlorides to form esters. Imagine reacting 12 alcohols and eight acid chlorides. Reacting each alcohol with each acid chloride gives a total of 12 x 8 = 96 esters. The additions are done by robots under computer control (see back page) and a particular ester occupies a particular place in the resulting array.

Q1. Write the equation for the reaction of an alcohol (ROH) with an acid chloride (R'COCl).

Q2. What ester will be present in the tube which is shaded in the Figure?



These use computer-controlled syringes, called robots, to carry out repetitive chemical techniques (such as adding specified quantities of reagents). Another approach is to carry out reactions on compounds while they are bonded to polymer beads – 'solid phase techniques'. The final product can be obtained in a pure form by washing off excess reagents from the polymer. The impetus to develop these techniques has been improvements in methods for rapidly screening large numbers of compounds for potential drug activity. (A vast library of compounds is of little use if it takes years to screen it.) Preliminary screening for many types of drug activity can now be done by measuring the ability of a compound to affect enzymes and bind to receptors on cells. This initial testing can be done in a test-tube (*in vitro)* rather than on living things (*in vivo)*. Only promising compounds are developed further.

•Solid phase reactions

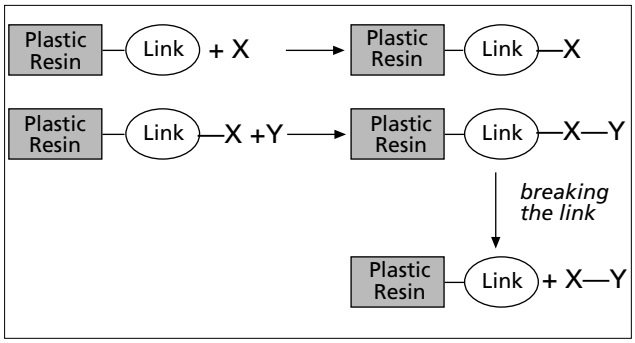


Figure 1. Solid phase chemistry

In solid phase organic chemistry, the starting material for a reaction is bonded covalently to beads of a plastic resin – often polystyrene-based – via a linking group. A reaction is done on the starting material, X, while it is attached to the resin. The product molecule remains bonded to the resin and can be released by breaking the link, or retained on the bead for further reaction (*Fig 1*). The chemistry of the link is important as it must be capable of being broken to release the product under conditions which will not break bonds in the new molecule which has been formed. This technique was originally developed by the American chemist Bruce Merrifield for making peptides (polyamino acids a few amino acids long) but the principle can be applied to all sorts of reactions. The *principle* of the technique is described below. (Some details have been omitted to keep things simple.)

Amino acids have the general formula $H_2NCH(R)CO_2H$ and can link together by the reaction of an $–NH_2$ group on one amino acid with the $–CO_2H$ on the next (*Fig 2*).

The process of making polypeptides

- The $–CO_2H$ group of the first amino acid, A, is bonded to a $–CH_2Cl$ group attached to the polystyrene resin. This is the link (*Fig 3*).
- The second amino acid has its $–NH_2$ group 'protected' – it is represented by HO–B–P, where P is the protecting group. This prevents this amino acid from reacting with other molecules of the same amino acid rather than the ones bonded to the resin.
- A solution of the second amino acid is then added to the resin and a dipeptide, A–B, is formed attached to the resin.
- The protecting group is then removed.
- A solution of a third amino acid (with its $–NH_2$ group protected) is added to the resin and a tripeptide, A–B–C, is formed, attached to the resin.
- The process of adding amino acids to the resin can be carried out up to 100 times to produce a polypeptide with different amino acids in any required sequence, attached to the resin.
- Finally the polypeptide can be released from the resin by breaking the link.

Q3. What molecule is eliminated as the link between the amino acid and the resin is formed?

Q4. Draw the dipeptide which could be formed in solution if the second amino acid (represented by $HO_2CCH(R')NH_2$) did not have its $–NH_2$ group protected.

Q5. What other peptides could also be formed *in solution* at the same time?

Q6. Look up the reactions of amines and of carboxylic acids. Use them to help you suggest how (a) the $–NH_2$, (b) the $–CO_2H$ group of an amino acid might be protected.

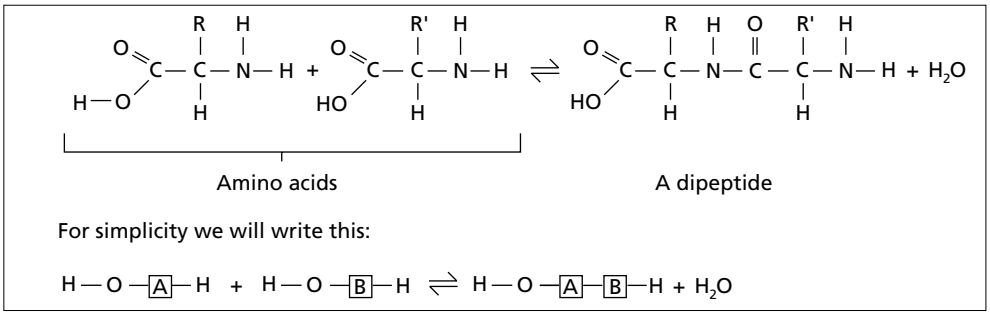
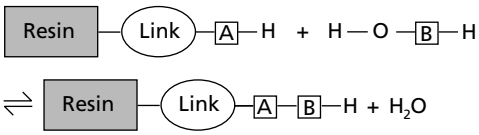


Figure 2. Forming a dipeptide

The advantages of this solid phase approach are:

- reactions forming peptides do not always go to completion. A large excess of the second (and subsequent) amino acids can be used to drive the equilibrium to the right (by Le Chatelier's principle):



- there is no need to extract the intermediate di- and tripeptides from the reaction mixture. The previous amino acid can simply be washed away while the product remains on the resin; and
- purification of the final product is easy. The resin beads can simply be filtered-off from the final reaction mixture and washed.

Q7. Look at the equation for making a dipeptide. How else could the equilibrium be forced to the right?

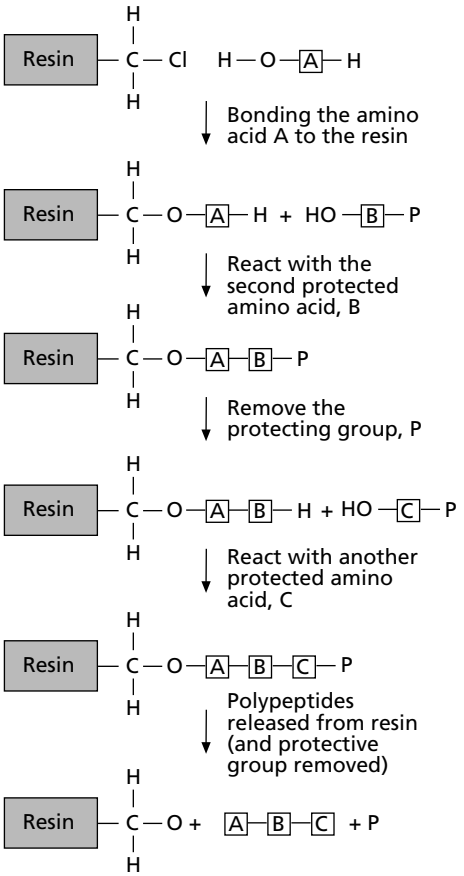
Split synthesis (Mix and divide)

This technique has the potential to make vast arrays of related compounds to start a chemical library and it is currently being developed. Imagine there are three monomers, A, B and C which react together like amino acids and which can be bonded to a resin (*Fig 4*).

1. We take a batch of resin beads, split it into three and react each portion with a different monomer in a different reaction vessel to give resin-A, resin-B and resin-C respectively.
2. All the beads are now mixed and again split into three portions which each contain $\frac{1}{3}$ resin-A, $\frac{1}{3}$ resin-B and $\frac{1}{3}$ resin-C.
3. Each portion is then reacted with A, B and C again. So vessel 1 will contain a mixture of resin-A-A, resin-B-A and resin-C-A. Vessel 2 will contain a mixture of resin-A-B, resin-B-B and resin-C-B and so on for vessel 3 so that we now have 9 compounds attached to beads.

Q8. What will vessel 3 contain?

4. The beads are mixed again, split into three once more and reacted with A, B and C. The resulting 27 compounds are shown in *Fig 4*. This is the beginning of a library of compounds. With more monomers and more steps we could have produced many more compounds. Three steps and three monomers gave 27 (3^3) compounds. The general rule is that with x monomers and n steps, we get x^n compounds. This soon leads to very large numbers indeed.



P = A group protecting the amino acid (to avoid the amino acid reacting with other molecules of amino acid in solution.)

Figure 3. Forming a polypeptide

Q9. How many compounds would you get with 20 starting materials (the number of naturally-occurring amino acids) and 4 steps?

Figure 4. Split synthesis

Stage	Reaction vessel 1 (A)	Reaction vessel 1 (B)	Reaction vessel 1 (C)	
1	Resin + A	Resin + B	Resin + C	3 compounds
		MIX		
2	Resin-A+A Resin-B+A Resin-C+A	Resin-A+B Resin-B+B Resin-C+B	?	9 compounds
		MIX		
3	Resin-A-A+A Resin-B-A+A Resin-C-A+A Resin-A-B+A Resin-B-B+A Resin-C-B+A Resin-A-C+A Resin-B-C+A Resin-C-C+A	Resin-A-A+B Resin-B-A+B Resin-C-A+B Resin-A-B+B Resin-B-B+B Resin-C-B+B Resin-A-C+B Resin-B-C+B Resin-C-C+B	Resin-A-A+C Resin-B-A+C Resin-C-A+C Resin-A-B+C Resin-B-B+C Resin-C-B+C Resin-A-C+C Resin-B-C+C Resin-C-C+C	27 compounds
		MIX		

RS•C

5. We now have 27 compounds attached to the beads. Any one individual bead has *only one compound* attached to it, although we do not know which particular compound this is. A gram of resin contains several million beads and, as in this example we have only 27 compounds, there will be very many beads containing each compound.

Identifying the compounds

The problem here is that each bead contains only a very small amount of compound – of the order of a few femtomoles (1 femtomole = 1 x 10^{-15} mol). A number of analytical techniques are used to determine the structure of compounds released from the beads. One of the most useful is mass spectrometry because it is sensitive enough to be usable with the very small amounts of substance obtained.

Q10. If a bead contains 6 femtomoles of compound, how many molecules is this?

• Screening for activity

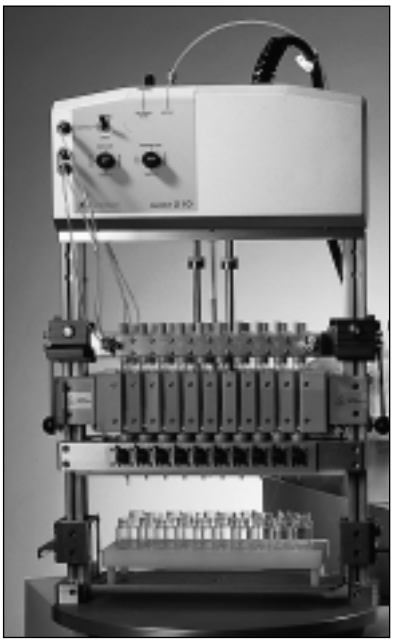
Present techniques of screening for biological activity require amounts of substance of the order of micromoles or nanomoles at the very least – much more than is present on a single bead. Chemists are working on methods of getting more product on a single bead as well as improving the sensitivity of the screening tests.

Q11. The development of combinatorial chemistry to synthesise potential drugs was driven by an advance in biology.

a) Describe this advance.

b) What two advances in *chemical* techniques were also required to make combinatorial chemistry useful?

The use of combinatorial chemistry to accelerate the rate of discovery of new chemical entities gives rise to issues in other areas. For example, the handling and storage of the analytical and biological data that are produced. Solving these problems is a challenge for chemists now and in the future.



Robot equipment for parallel synthesis produced by Argonaut Technologies

CHIPS IN A TEABAG Box 3

One approach to the problem of identifiying compounds is tagging beads so that the tag records the history of each bead – which reagents it has been treated with and in what order. This can be done chemically, or with a microchip which emits radiowaves. In this case the tag is placed in a capsule, rather like a teabag, which also contains a quantity of resin. The microchip can record the history of the batch of beads in the capsule.

Even without this method, we do know something about the identity of the compounds that we have produced. In the example above, we know that after the third stage, all the beads in vessel three have C on their free end.

Bruce Merrifield US pioneer of solid phase chemistry (photograph by courtesy of Ingbert Grüttner, The Rockefeller University)

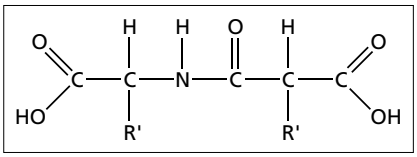
Acknowledgement

The Royal Society of Chemistry and the author, Ted Lister, thank Dr Mike Davies of Celltech Therapeutics for his help in preparing this article.

RS•C

Answers...

•Combinatorial chemistry

1. ROH + R'COCl → R'COOR + HCl
2. Ethyl propanoate
3. HCl
4. 

5. The corresponding tripeptide, tetrapeptide *etc.*
6. Based on A-level chemistry, students might suggest the following:

a) Converting the $-CO_2H$ group into an ester

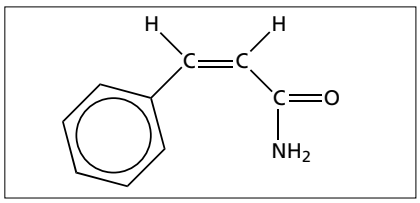
b) Converting the $-NH_2$ group into an N-substituted amine or an amide

However, neither of the compounds in (b) is really suitable as neither is easily removable.

7. Remove water as it is formed
8. Resin-A–C, resin–B–C and resin–C–C
9. 20^4 (160 000)
10. 3.6×10^9
11. a) High throughput screening for biological activity.
    b) Robotic techniques for synthesis. and sensitive analytical techniques

•Computational chemistry

1. React cinnamic acid with ammonia and then heat to dehydrate the ammonium salt produced.
2. Derivatives might include: ester, acid chloride, anhydride, salts and the parent alcohol, addition products of the double bond *etc.*
3. The $-CONH_2$ group can hydrogen bond with water but the rest of the molecule is non-polar so the solubility is not great.
4. $PhCHCHCONH_2 + H_2O \rightarrow PhCHCHCOOH + NH_3$

   A nucleophilic substitution reaction.
5. a ~120°, b ~105°, c ~121°, d = 120°
6. a) This is a benzene ring and undergoes electrophilic substitution reactions.
   b) This is an alkene and undergoes electrophilic addition reactions.
   c) This is a carboxylic acid and undergoes acid-base reactions and nucleophilic substitution reactions
7. *trans*-3-phenylprop-2-enamide.
8. 

9. The C-$CONH_2$ bond, the C–N bond and the C–Ph bonds can rotate, the others cannot.