MANCHESTER
1824

The University
of Manchester

http://utopiadocs.com/

@utopiadocs @srp

MANCHESTER 1824

The University of Manchester

Not a Hamburger

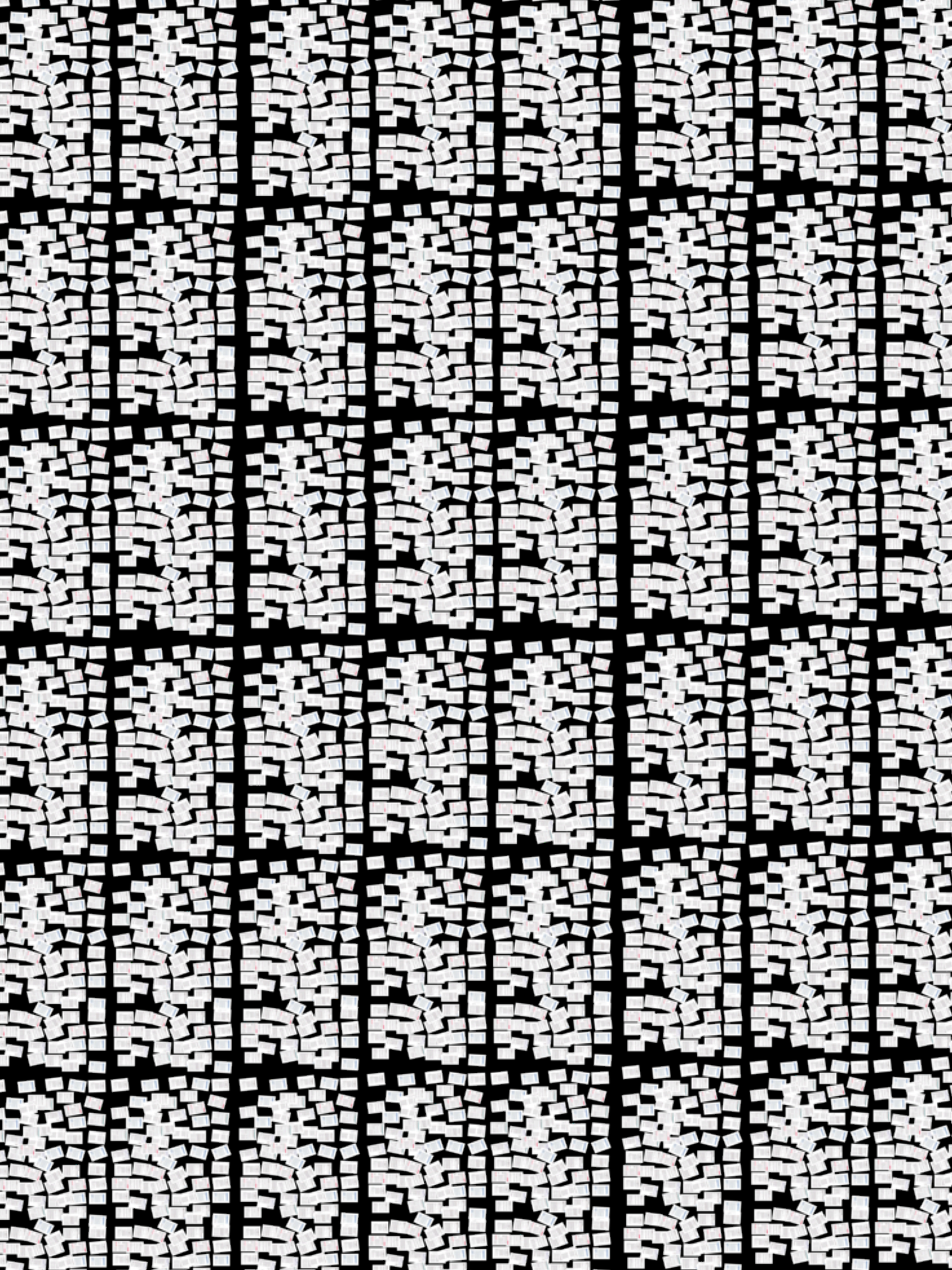
The treachery of representations

Steve Pettifer
The University of Manchester

http://utopiadocs.com/
@utopiadocs @srp

The famous pipe. How people reproached me for it! And yet, could you stuff my pipe? No, it's just a representation, is it not? So if I had written on my picture "This is a pipe," I'd have been lying!

René Magritte (1928)



Bairoch's Lament

"It is quite depressive to think that we are spending millions in grants for people to perform experiments, produce new knowledge, hide this knowledge in often badly written text and then spend some more millions trying to second guess what the authors really did and found"

Bairoch A (2009) The future of annotation/biocuration. Nature Precedings doi:10.1038/npre.2009.3092.1.

The Problem?

I make a thing

You find a thing

You have a thing

I make a thing
I send the thing to a publisher
Publisher gets some people to look at the thing
People don’t completely like the thing
I change my thing
Now some people like my thing
The publisher makes my thing a bit nicer to look at

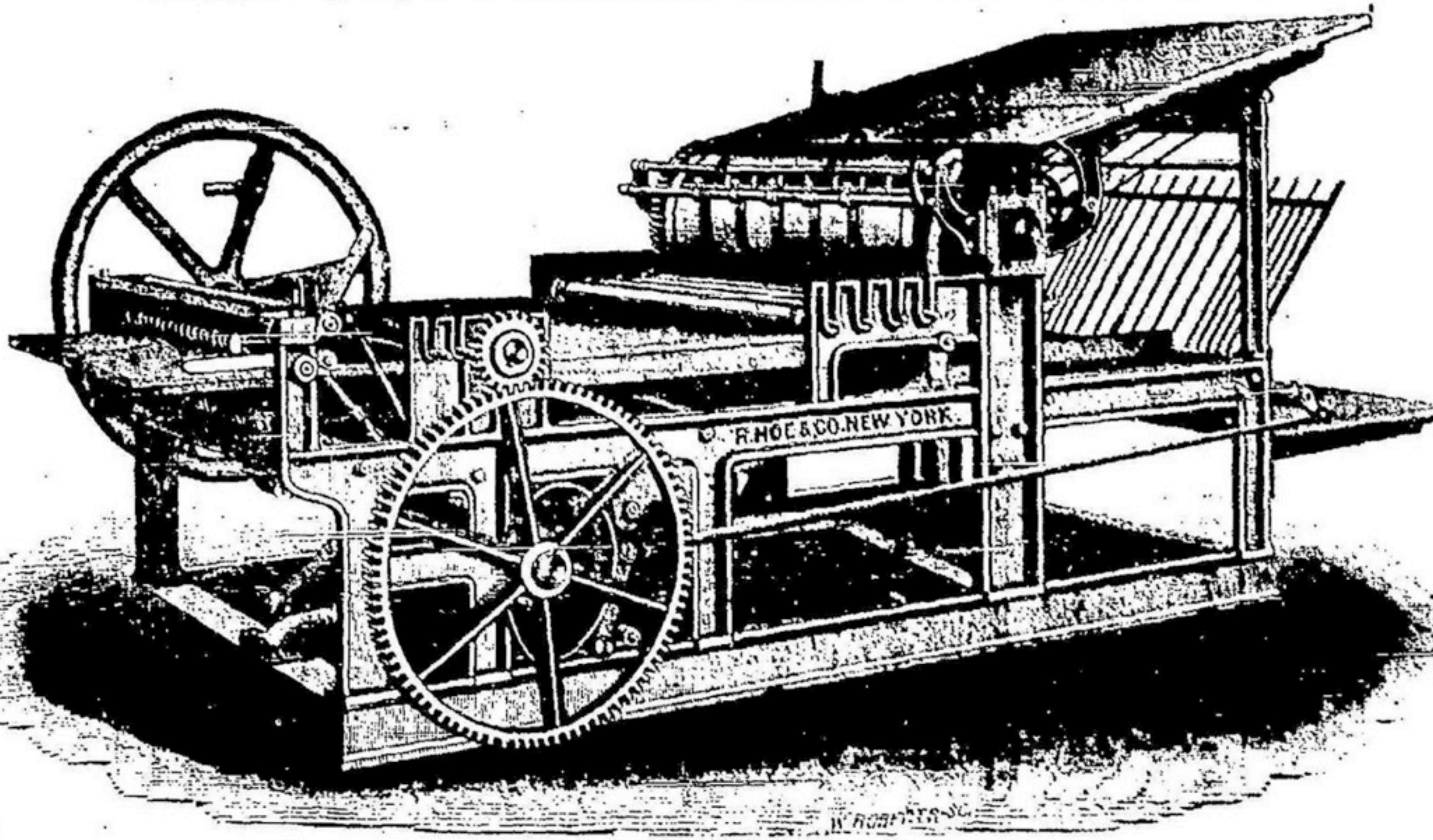
and puts its in a special place
and you can find a thing

and read it

Publish or Perish:

The Prisoner's Dilemma

HAND STOP-CYLINDER PRINTING MACHINE.



HAND STOP-CYLINDER PRINTING MACHINE.

The Deadly Embrace

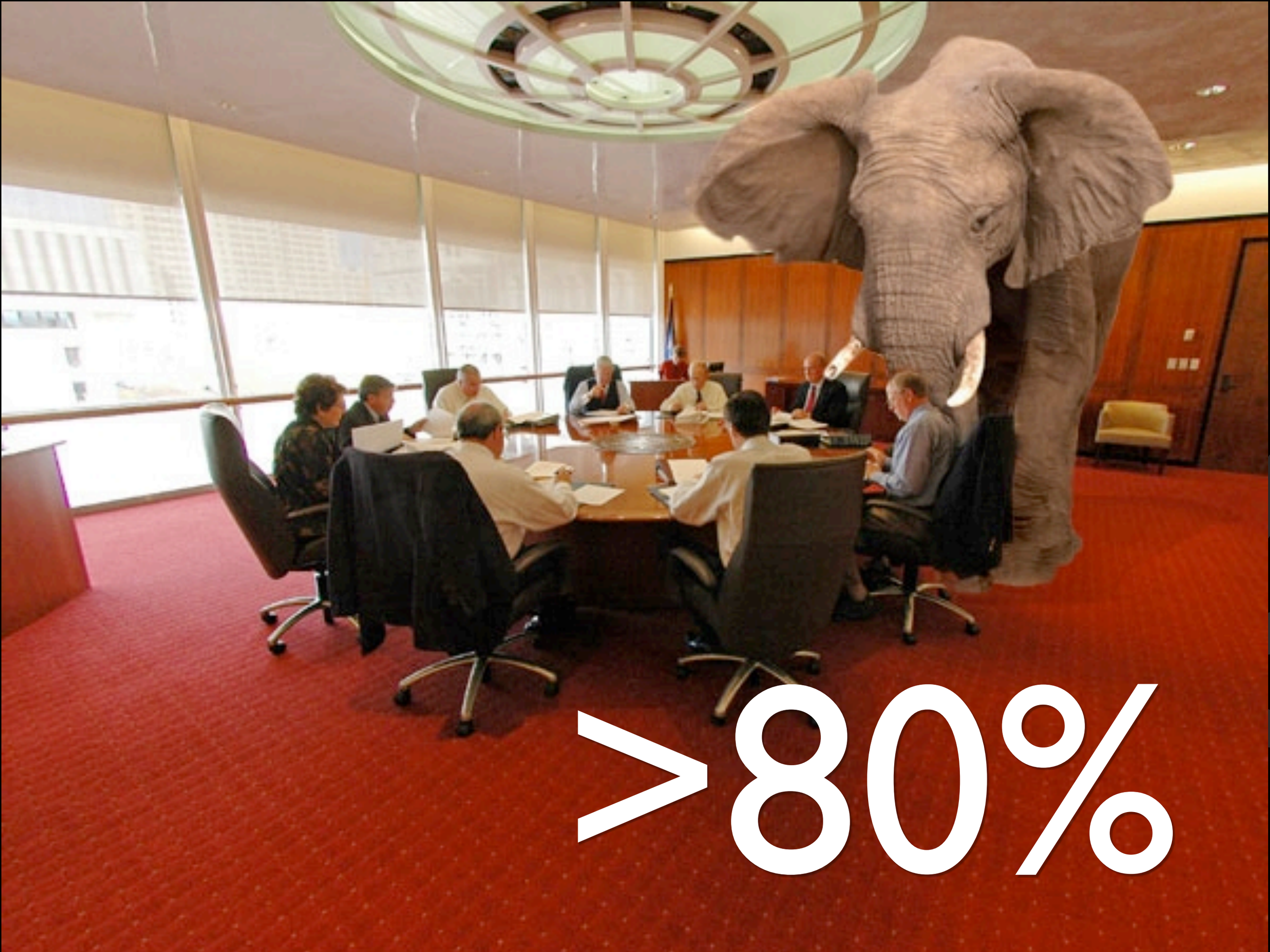
PDF
Adobe

“An insult to science”

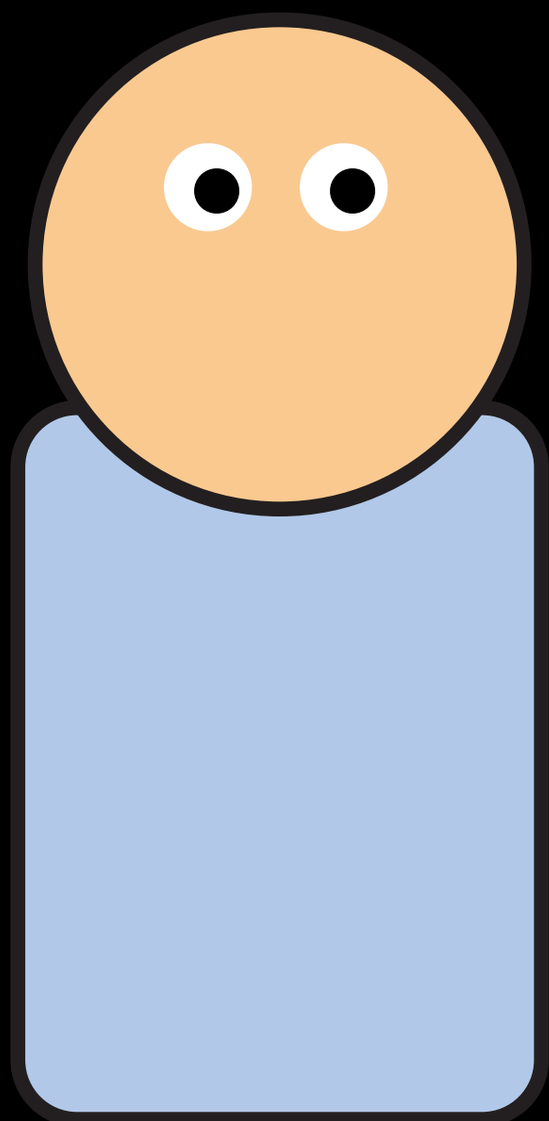
“Like inventing the telephone and then using it to transmit Morse Code”

“Antithetical to the spirit of the Web”

“A hamburger that we are trying to turn back into a cow”

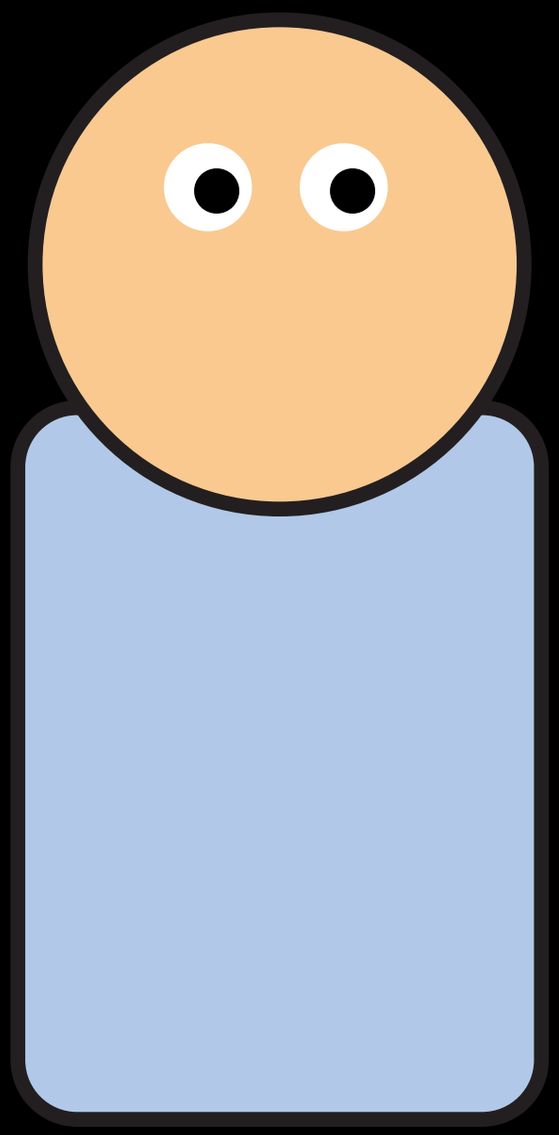

>80%

I've got some things, and I need them ordered according to their size, smallest first, biggest last.

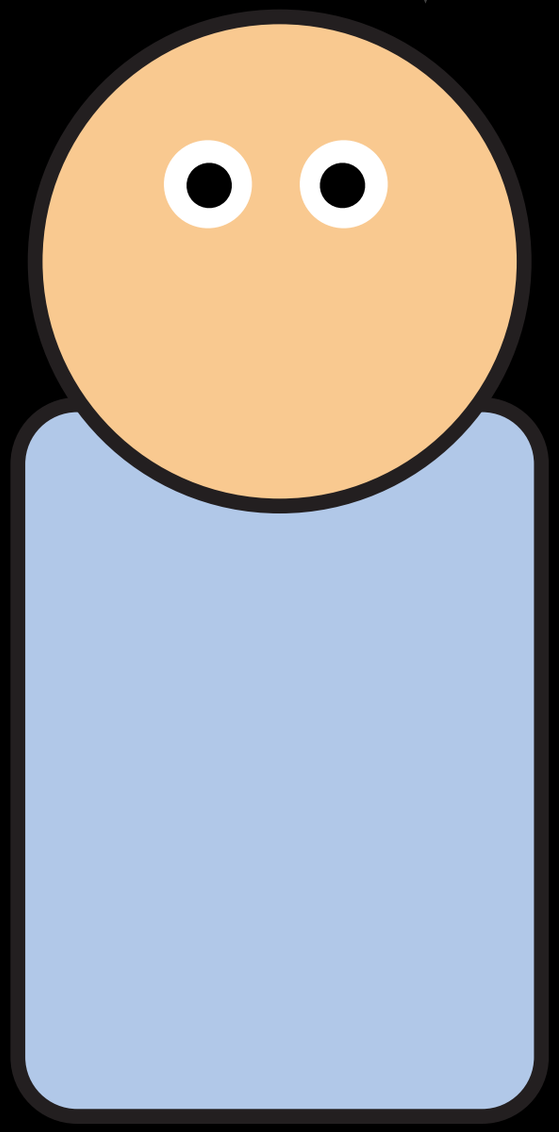


So you mean...

$$\forall i, j \in dom(s).i \leq j \Rightarrow s[i] \leq s[j]$$



I've no idea.
I just want some things,
smallest first, biggest last

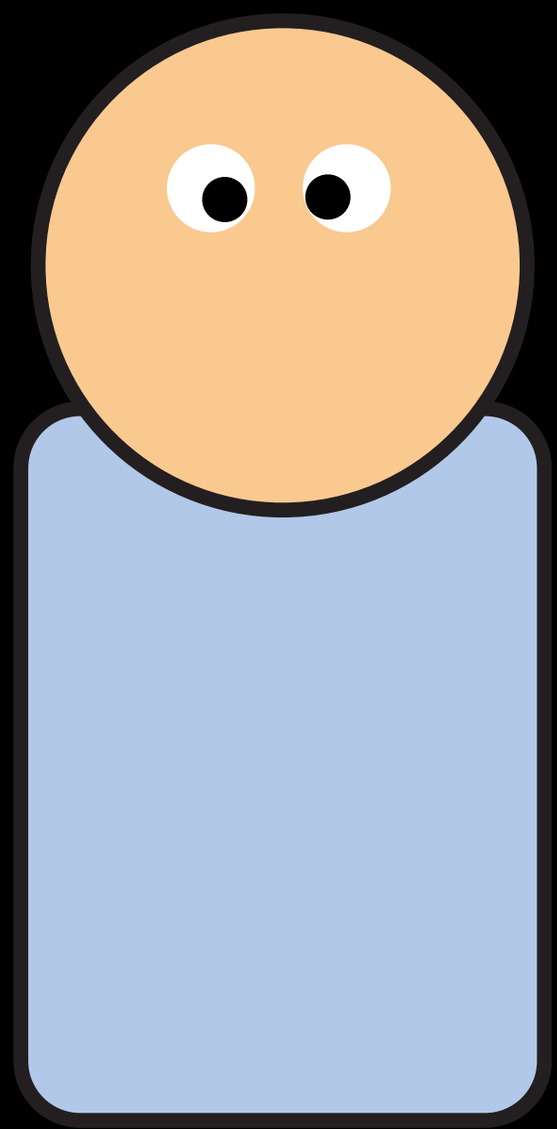


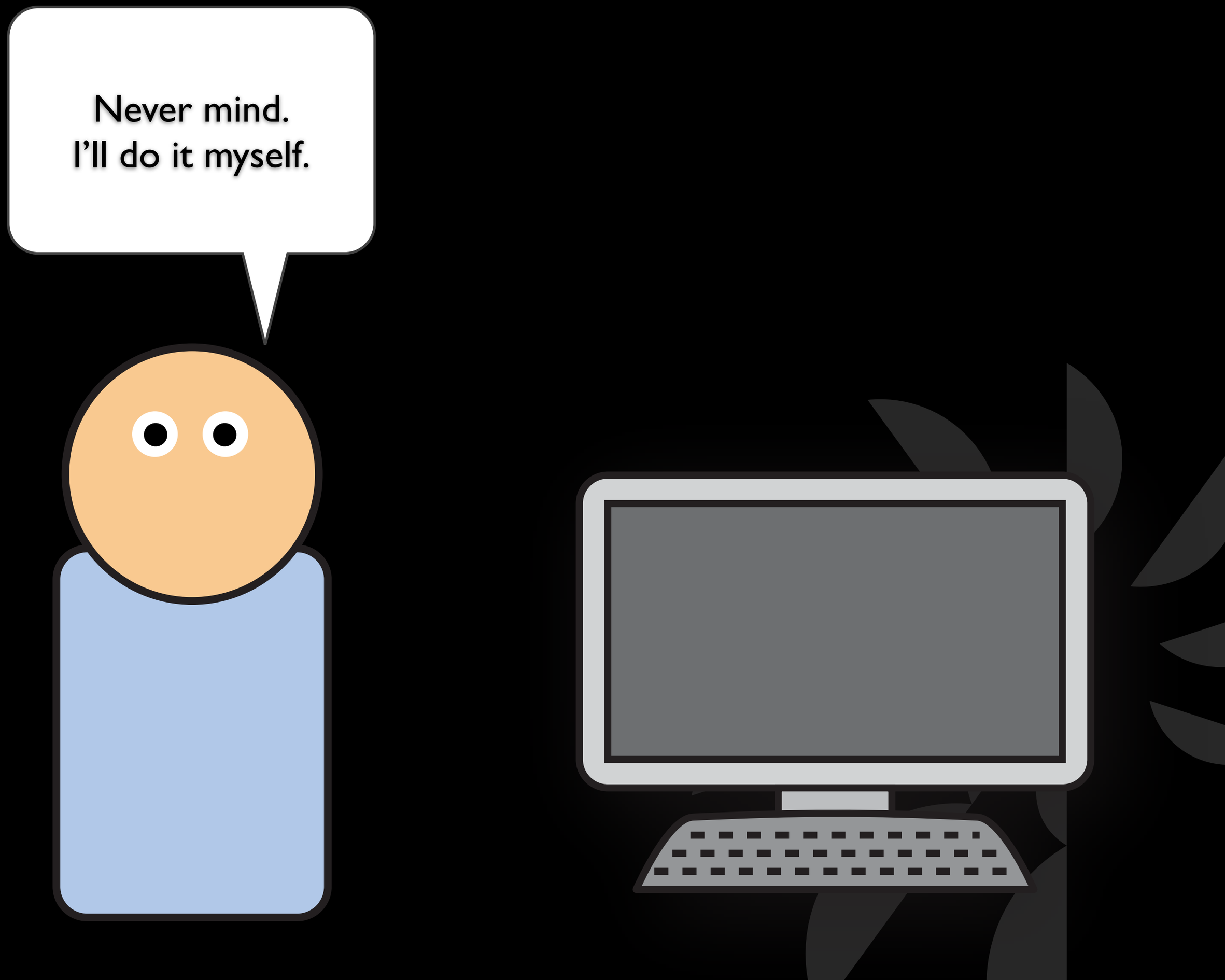
```
def main():
   array = [1, 7, 4, 9, 4, 7, 2, 3, 0, 8]
   print("Before sort:-")
   print(array)
   bubbleSort(array)
   print("After sort:-")
   print(array)

def bubbleSort(array):
   swapHappened = True

   while swapHappened:
      swapHappened = False

      for x in range(0, len(array)-1):
         if array[x] > array[x+1]:
            # Swap data
            array[x], array[x+1] = array[x+1], array[x]
            swapHappened = True
```

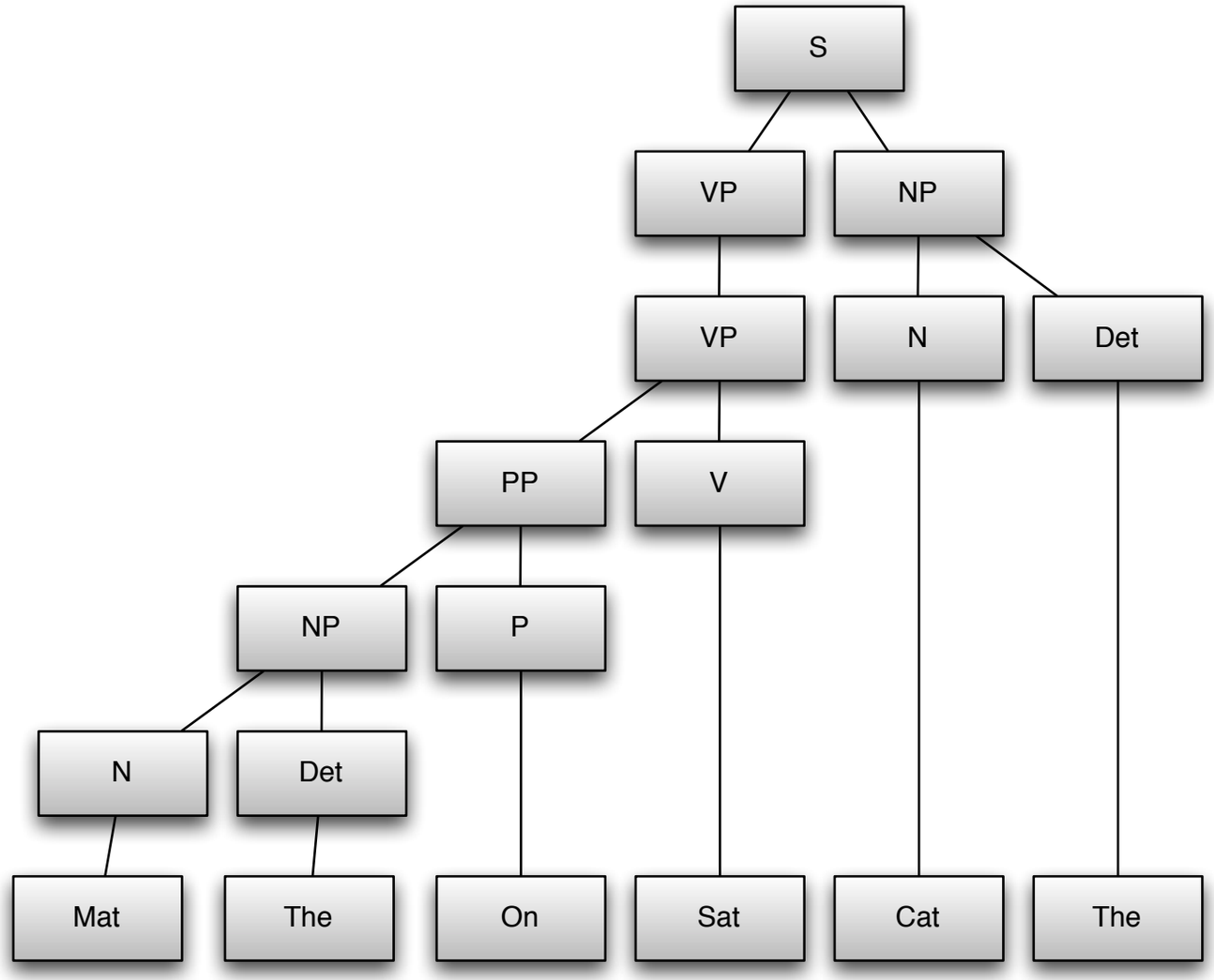



Never mind.
I'll do it myself.

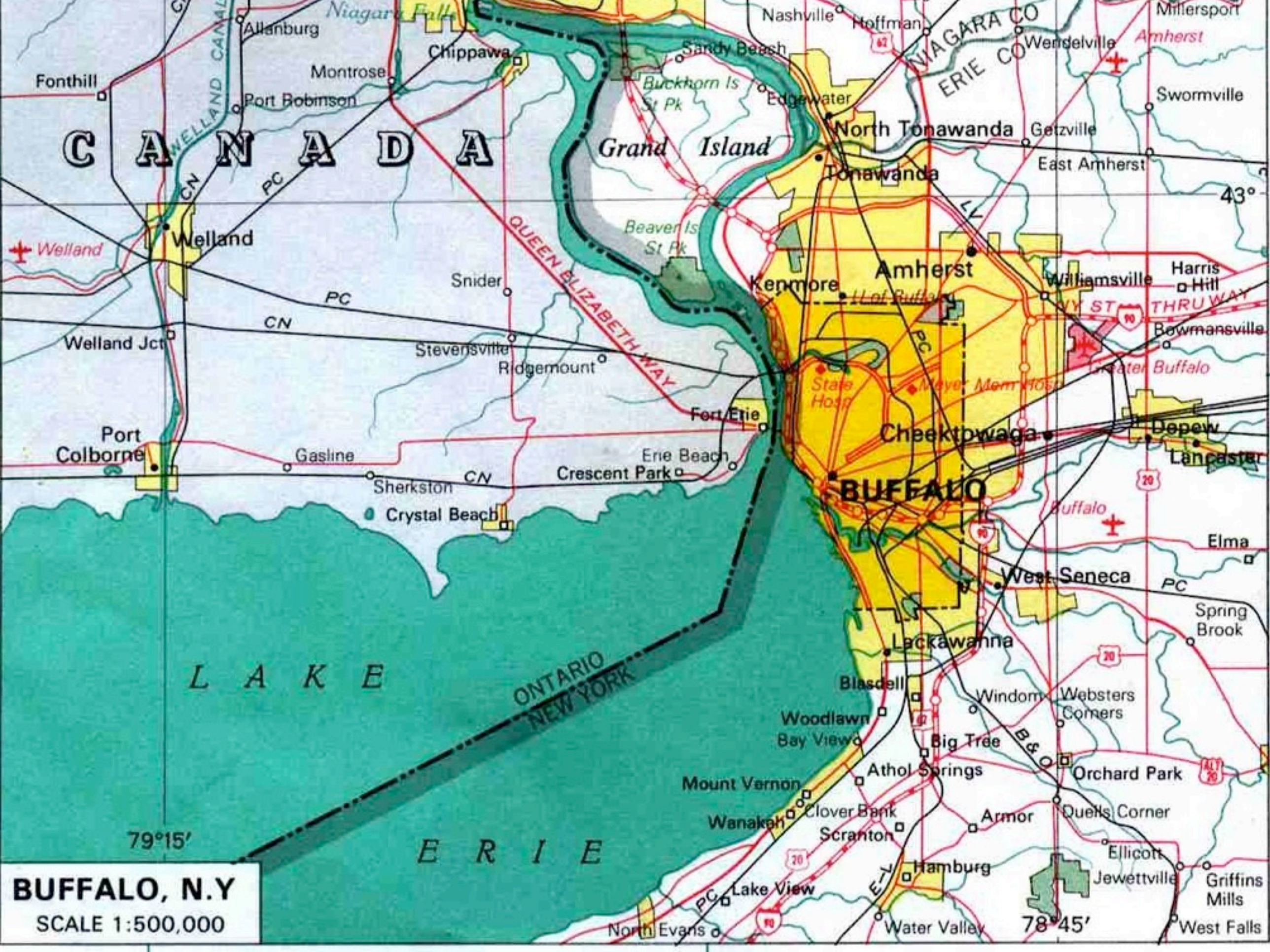
Mr Tiddles was resting
in the CozeeCat Bag

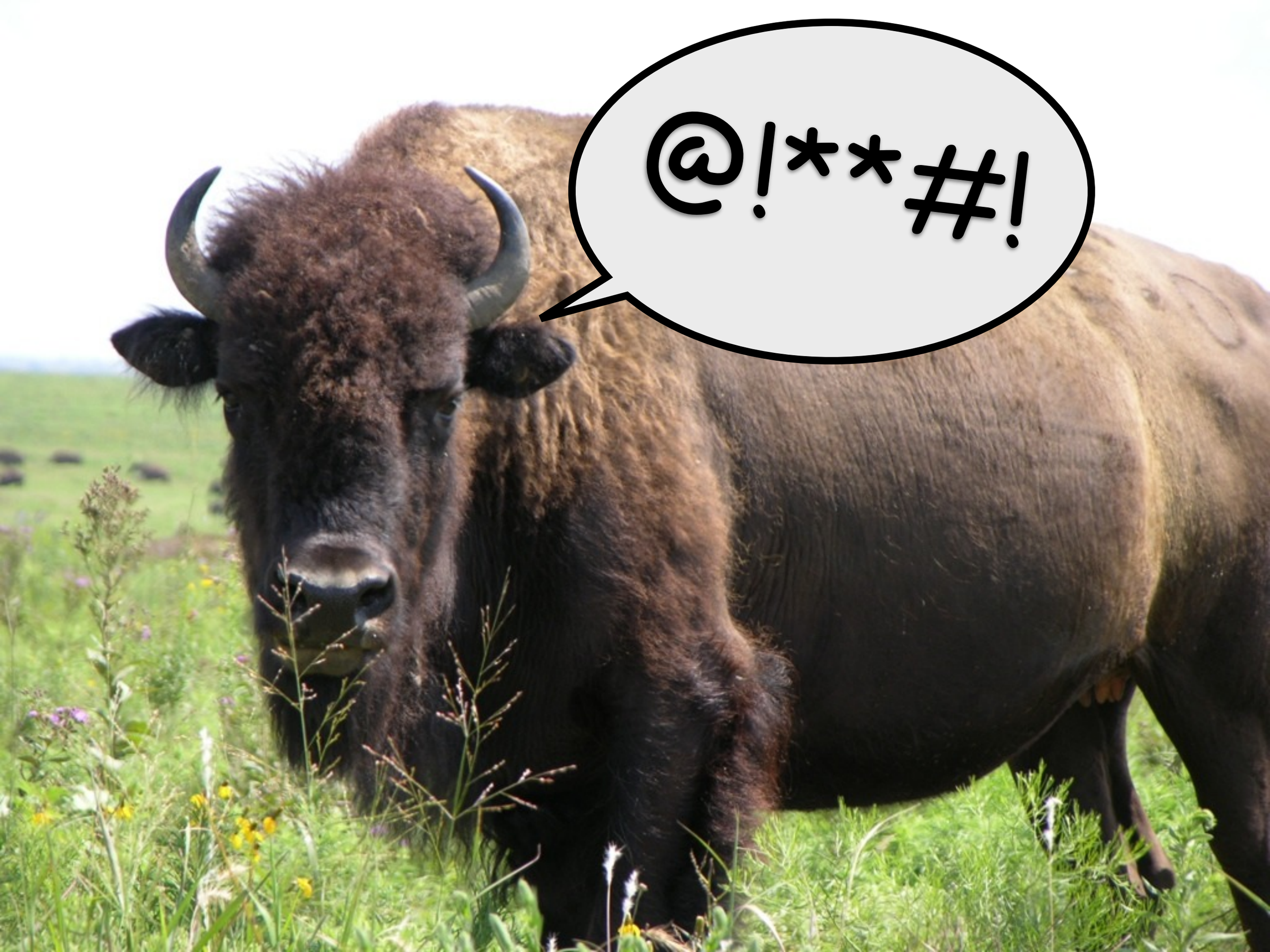
Mr Tiddles was resting in the CozeeCat Bag

The cat sat on the mat



Buffalo buffalo
Buffalo buffalo buffalo
buffalo Buffalo buffalo


CANADA
LAKE ERIE
ONTARIO
NEW YORK
Grand Island
Niagara Falls
Allanburg
Chippawa
Montrose
Port Robinson
Fonthill
WELLAND CANAL
Welland
Welland Jct
Port Colborne
Gasline
Sherkston
Crystal Beach
Crescent Park
Erie Beach
Fort Erie
Snider
Stevensville
Ridgemount
QUEEN ELIZABETH WAY
Sandy Beach
Buckhorn Is St Pk
Beaver Is St Pk
Nashville
Hoffman
Edgewater
North Tonawanda
Tonawanda
NIAGARA CO
ERIE CO
Wendelville
Millersport
Amherst
Swormville
Getzville
East Amherst
Kenmore
Amherst
Williamsville
Harris Hill
Bowmansville
Greater Buffalo
State Hosp
Meyer Mem Hosp
Cheektowaga
Depew
Lancaster
BUFFALO
Buffalo
Elma
West Seneca
Spring Brook
Lackawanna
Blasdell
Woodlawn
Bay View
Windom
Websters Corners
Big Tree
Athol Springs
Orchard Park
Mount Vernon
Clover Bank
Wanakah
Scranton
Armor
Duells Corner
Ellicott
Hamburg
Jewettville
Griffins Mills
Lake View
North Evans
Water Valley
West Falls
43°
79°15′
78°45′
BUFFALO, N.Y
SCALE 1:500,000



Buffalo buffalo

Buffalo buffalo buffalo

buffalo Buffalo buffalo

Buffalo bison (whom)

Buffalo bison bully, also

bully Buffalo bison

“Methods”

“Results and Conclusions”

Comprehensible to machine-ness

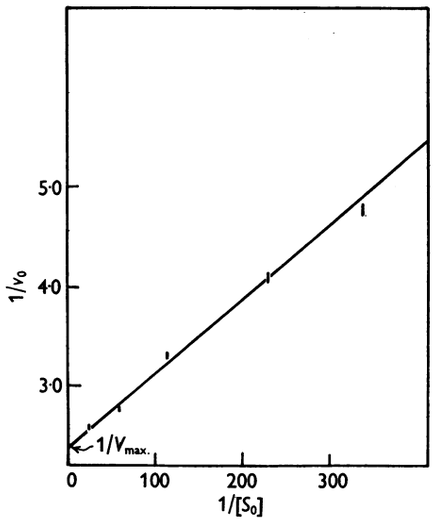


Fig. 4. Determination of K_m of benzoyl-L-arginineamide at pH 7·60 in 0·01 M phosphate buffer, 0·2 M-NaCl and 2% (w/v) formaldehyde. Each reaction mixture contained 3·3 mg. of enzyme. $[S_0]$ and v_0 are the respective initial substrate concentrations and rates. K_m was calculated from the slope by means of the equation

$$\frac{(V_{max.})_{[S]\to\infty}}{v_0} = \frac{K_m}{[S_0]} + 1.$$

(Neurath & Schwert, 1950). It would be unlikely, therefore, that a free amine group (ammonium ion at pH 7·6) would be found in the region of the active site. It is of interest to note that acetylation of trypsin results in no loss of esterase activity (Jansen, Nutting, Jang & Balls, 1949).

SUMMARY

1. Although formaldehyde alters the chemical structure of trypsin it has no effect on the activity of the enzyme at the pH optimum.

2. A method for following the trypsin-catalysed hydrolysis of amides based on the above observation has been described.

3. The enzyme–substrate dissociation constant for benzoyl-L-arginineamide has been determined ($K_m = 3·1 \times 10^{-3}$ M) as well as the first-order specific rate of decomposition of the enzyme–substrate complex ($k_3 = 0·043$ sec.$^{-1}$).

4. The pronounced deceleration of trypsin-catalysed hydrolysis with time has been explained in terms of the self-digestion of the enzyme (a process which is inhibited by the substrate).

The author wishes to express his appreciation to Dr Herbert Gutfreund for many helpful discussions and to Professor F. J. W. Roughton, F.R.S., for making available the facilities of the Department of Colloid Science, Cambridge.

This work was done wholly under a fellowship from the American Cancer Society upon recommendation of the Committee on Growth of the National Research Council (U.S.A.).

REFERENCES

Bergmann, M., Fruton, J. S. & Pollok, H. (1939). *J. biol. Chem.* **127**, 643.
Gutfreund, H. (1954). *Trans. Faraday Soc.* **50**, 624.
Harmon, K. M. & Niemann, C. (1949). *J. biol. Chem.* **178**, 743.
Huang, H. T. & Niemann, C. (1951). *J. Amer. chem. Soc.* **73**, 475.
Iselin, B. M. & Niemann, C. (1950). *J. biol. Chem.* **183**, 403.
Jansen, E. F., Nutting, M.-D. F., Jang, R. & Balls, A. K. (1949). *J. biol. Chem.* **179**, 189.
Lineweaver, H. & Burk, D. (1934). *J. Amer. chem. Soc.* **56**, 658.
Neurath, H. & Schwert, G. W. (1950). *Chem. Rev.* **46**, 69.
Northrop, J. H., Kunitz, M. & Herriott, R. M. (1948). *Crystalline Enzymes.* Columbia University Press.
Schwert, G. W. & Eisenberg, M. A. (1949). *J. biol. Chem.* **179**, 665.
Schwert, G. W., Neurath, H., Kaufman, S. & Snoke, J. E. (1948). *J. biol. Chem.* **172**, 221.

The Amide Groups of Insulin

By F. SANGER,* E. O. P. THOMPSON† and RUTH KITAI
Department of Biochemistry, University of Cambridge

(*Received* 6 *September* 1954)

Previous work has shown that insulin consists of two types of polypeptide chains the amino acid sequences of which have been determined by partial hydrolysis experiments (Sanger & Tuppy, 1951*a*, *b*; Sanger & Thompson, 1953*a*, *b*). For an assumed molecular weight of approximately 6000, which is the simplest chemical unit consisting of one phenylalanyl and one glycyl chain, there are six amide groups present (Rees, 1946). From a quantitative study of the products of reduction of insulin ester with lithium borohydride Chibnall & Rees (1952) showed that the three aspartic acids, one of which is the *C*-terminal residue of the glycyl chain, are all present as amides, whereas three glutamic acid residues are amidized and the other four have free carboxyl groups. The following paper describes experiments to locate these amide groups on the individual dicarboxylic acid residues. To achieve this, use has been made of the fact that amide groups are split rapidly by the action of acid, but are not in general attacked by proteolytic enzymes. Thus peptides obtained from enzymic digests contain the original amide groups intact, whereas those obtained from acid hydrolysates do not. Two methods are thus available for identifying amide groups. In the first, the ionophoretic rates of the peptides from enzymic digests are compared with the rates of other similar peptides obtained from acid and enzymic hydrolysates. From these results it is possible to calculate the number of charged groups on the peptides and hence the number of carboxyl groups that are masked as amides. In the second method the number of amide groups present in the peptides from enzymic digests is determined directly by an estimation of the ammonia produced by hydrolysis with acid.

A number of peptides were identified from peptic and chymotryptic hydrolysates of fraction *A* (Sanger & Thompson, 1953*b*) and these were used to locate most of the amide groups. However, other peptides were required, especially those containing only one of the two glutamic acid residues in positions 4 and 5 (see Table 7), and so hydrolysates obtained by the action of two other enzymes (papain and mould protease) were investigated and the results are reported. The mould protease used was first isolated in a crystalline form by Crewther & Lennox (1950), but was not considered to be a pure substance. It is extremely active though rather unspecific in its attack.

The abbreviations and reference numbers to the peptides are those reported in previous papers (Sanger & Tuppy, 1951*a*, *b*; Sanger & Thompson, 1953*a*, *b*). Pyr is used to denote a residue of pyrrolidone carboxylic acid (5-oxopyrrolidine-2-carboxylic acid).

EXPERIMENTAL

Action of mould protease on fraction A (*Expt. A*m). The mould protease used was a crystalline sample obtained through Dr R. R. Porter of the National Institute for Medical Research, Mill Hill, London.

A solution of fraction *A* (30 mg.) in 5 ml. water was brought to pH 8 and 1·5 mg. mould protease were added. After incubating at 37° for 24 hr. the mixture was boiled to inactivate the enzyme and taken to dryness *in vacuo*. Samples (5 or 10 mg.), after centrifuging to remove insoluble matter, were fractionated on Whatman no. 4 or no. 3 paper chromatograms, using phenol–0·3% NH_3 followed by butanol–acetic acid as solvents and the peptide spots eluted and subjected to hydrolysis and end-group determination as described in previous papers (Sanger & Tuppy, 1951*a*, *b*; Sanger & Thompson, 1953*a*, *b*).

Action of papain on fraction A (*Expt. A*pa). The papain used was a commercial preparation supplied by British Drug Houses Ltd. A sample (16 mg.) was activated before use by incubation for 1 hr. with 1·6 ml. 1% NaCN solution at 37°. The activated enzyme (0·5 ml.) was then added to a solution of 50 mg. fraction *A* in 2 ml. water. The pH was adjusted to 7 with dilute NH_3 and the mixture incubated at 37° for 24 hr. Subsequent working up was as described for mould protease.

*Preparation of peptide A*p5 *by ionophoresis.* Peptide *A*p5, which consists of the first thirteen residues of the glycyl chain (Sanger & Thompson, 1953*b*) was prepared from insulin by making use of its high mobility as an acid in 0·2 N acetic acid, owing to its content of three cysteic acid residues.

Insulin (50 mg.) in 10 ml. 0·01 N-HCl was treated with 1·0 mg. pepsin and incubated at 37° for 24 hr. After being taken to dryness, the residue was oxidized with performic acid (Sanger, 1949*a*). The formic acid was removed *in vacuo*, the residue dissolved in 1·5 ml. water and applied along the centre line of a sheet of Whatman no. 3 filter paper. It was then subjected to ionophoresis in an apparatus similar to that of Durrum (1950). A potential of 220 v was applied for 21 hr. A strip was cut from the paper and tested with ninhydrin. The *A*p5 was present in a band which stained rather weakly with ninhydrin and had moved 12·5 cm. towards the anode. From a ninhydrin determination on a hydrolysate of the eluted material it appeared that the yield of *A*p5 was about 30% of the theoretical yield from insulin. In several experiments a faint slower moving band could be distinguished, which appears to be peptide *A*p9 (the first ten residues of the glycyl chain). For further purification the *A*p5 was subjected to ionophoresis at pH 3·7 in the apparatus of Michl (1951). The material obtained from 50 mg. insulin was put on a 18 cm.-wide strip of no. 3 filter paper; the ionophoresis was continued for 3·5 hr. at 1500 v. The main *A*p5 band had moved 19 cm. towards the anode. Several fainter slow moving bands were also present.

Ionophoresis of peptides. The ionophoretic rates of peptides have been compared using the apparatus of Durrum (1950). The electrolyte used was 0·05 M ammonium acetate (pH 6·8). The peptides obtained from paper chromatograms were applied to a sheet of Whatman no. 1 filter paper either directly by the 'strip transfer' method, or after elution from the paper and drying in a desiccator. The potential used was 220 v; since the rates were not reproducible from experiment to experiment, only those done on the same paper were compared. If the original peptide spot had been treated with ninhydrin, the purple colouring matter usually migrated slightly towards the anode and was distinguished from the colour given by neutral spots on subsequent spraying, since they moved towards the cathode with the electroendosmotic flow.

Micro amide determinations on peptides. Estimations of amide N on the peptides eluted from paper chromatograms were carried out by the standard procedure (see Bailey, 1937), the ammonia being estimated by the micro-distillation method of Conway (1950). In order to determine the amount of peptide present in the samples, portions were subjected to hydrolysis, and the total amino acids estimated by the ninhydrin method of Moore & Stein (1948).

* Member of the scientific staff of the Medical Research Council.

† Present address: Wool Textile Research Laboratories, 343 Royal Parade Parkville, Victoria, Australia.

toward the anode, suggesting the glutamic acid residue is not present as an amide. The yellow colour with ninhydrin suggests that serine is *N*-terminal. It is tentatively suggested that synthesis may have occurred under the influence of the papain, but it is hoped to investigate this effect further.

Action of papain on Ap5. Peptide *Ap5* purified by a single ionophoresis, was hydrolysed with cyanide-activated papain as described above and the hydrolysate subjected to ionophoresis in 0·2M acetic acid overnight at 220v. Test strips were coloured with ninhydrin, and by the chlorination procedure of Rydon & Smith (1952). The distribution of the bands is shown in Fig. 4, and their composition in Table 4. Since band 5 gave no colour with ninhydrin but a strong colour by the chlorination method and contained glutamic acid, it is presumed to be a pyrrolidonoyl peptide. The presence of free tyrosine (band 2) is probably due to some impurity in the *Ap5*.

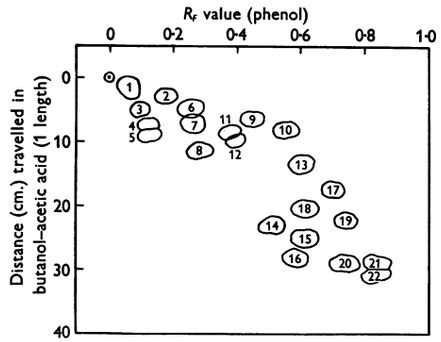


Fig. 2. Chromatogram of mould protease hydrolysate of fraction *A* (Expt. *Am*) (see Table 1).

Total amides of fraction A. Sanger (1949*a*) obtained a value of 15·15% for the amide-N of fraction *A* as percentage of the total N. This approximates to four residues per chain (theoretical 16·0%). Since it was essential to be certain of this value, it has been confirmed in the present work.

Bound NH_3 was removed before amide estimation by dissolving the fraction *A* (15–20 mg.) in 1 ml. 0·1M-K_2CO_3 and leaving in a vacuum desiccator over H_2SO_4 for 4 hr. The value obtained was 16·7% amide-N as percentage of total N. The theoretical value for five residues would be 19·2%.

Further evidence that there are four amide groups in fraction *A* was obtained from plotting a titration curve. Since fraction *A* contains no histidine, the only groups titrating in the range pH 2·5–7 are the

Table 1. *Peptides from mould protease hydrolysates of fraction A* (*Am*)

Spot no. (Fig. 2)	Approx. strength of peptide	Composition*	Probable structure
1	× × ×	Glu.[$CySO_3H$, Ser, Ala, Val]	Glu.$CySO_3H$.$CySO_3H$.Ala.Ser.Val.$CySO_3H$
2	× × ×	[$CySO_3H$, Ser, Ala, Val]	$CySO_3H$.Ala.Ser.Val.$CySO_3H$
3	×	[$CySO_3H$, Glu, Ser, Ala]	Mixture
4	× ×	[$CySO_3H$, Asp, Tyr]	Tyr.$CySO_3H$.Asp
5	×	[$CySO_3H$, Asp, Tyr]	—
6	× × × ×	[$CySO_3H$, Asp, Tyr]	Asp.Tyr.$CySO_3H$.Asp
7	×	[$CySO_3H$, Asp, Tyr]	—
8	×	—	Glutamic acid
9	× ×	[Asp]	Asparagine
10	× ×	[Glu]	Glutamine
11	{ × × ×	[Ser]	Serine
	×	[$CySO_3H$, Val]	Val.$CySO_3H$ and/or Ser.Val.$CySO_3H$
12	× ×	[$CySO_3H$, Glu, Gly, Val, Leu]	Gly.Ileu.Val.Glu.Glu.$CySO_3H$
13	×	—	Alanine
14	× × ×	Val.Glu	Val.Glu
15	× ×	Glu.[Glu, Leu]	Glu.Leu.Glu
16	× × ×	Leu.Glu	Leu.Glu
17	× ×	[Glu, Tyr]	Tyr.Glu
18	× ×	—	Tyrosine
19	×	[Glu, Val]	?
20	× ×	[Glu, Gly, Val, Leu]	Gly.Ileu.Val.Glu
21	× × ×	—	Leucine
22	× × ×	[Ser, Tyr, Leu]	Ser.Leu.Tyr

* Leucine and isoleucine were not distinguished. Both are recorded as Leu.

Spot 1. Faint colour with ninhydrin. *Spots* 1 *and* 2. May have an extra serine in the *C*-terminal position. *Spot* 3. Contained both glutamic acid and cysteic acid as *N*-terminal residues, so it was clearly a mixture. *Spot* 4. Grey colour with ninhydrin similar to tyrosine, suggesting tyrosine is *N*-terminal. *Spot* 6. Similar R_F value and colour with ninhydrin (yellow) to *Ap3* and *Apa3* (Asp.Tyr.$CySO_3H$.Asp). *Spots* 5 *and* 7. May be deamidated forms of *Am4* and *Am6* or one could be Asp.Tyr.$CySO_3H$. *Spot* 9. Yellow-brown with ninhydrin. *Spots* 12 *and* 22. Yellow with ninhydrin. *Spot* 20. On the chromatogram was a mixture with a tyrosine-containing peptide. It was purified by ionophoresis, yellow colour with ninhydrin. Same R_F value as peptide *A*1η3 (Gly.Ileu.Val.Glu).

Table 2. *Peptides from papain hydrolysate of fraction A* (*Apa*)

Spot no. (Fig. 3)	Approx. strength of peptide	Composition*	Probable structure
1	× × ×	[$CySO_3H$, Glu, Ser, Ala, Val]	Pyr.$CySO_3H$.$CySO_3H$.Ala.Ser.Val.$CySO_3H$
2	× ×	[$CySO_3H$, Asp, Glu, Tyr]	Glu.Asp.Tyr.$CySO_3H$.Asp
3	× ×	[$CySO_3H$, Asp, Tyr]	Asp.Tyr.$CySO_3H$.Asp
4	×	[$CySO_3H$, Asp, Tyr]	—
5	× × ×	[$CySO_3H$, Asp, Glu, Tyr, Leu]	Leu.Glu.Asp.Tyr.$CySO_3H$.Asp
6	×	[$CySO_3H$, Asp, Glu, Tyr, Leu]	—
7	×	—	Glutamic acid
8	× ×	—	Glycine
9	×	[Glu, Ser, Tyr]	?
10	× ×	—	Tyrosine
11	× × ×	[Glu, Tyr]	Tyr.Glu
12	×	[Glu, Tyr, Leu]	Tyr.Glu.Leu.Glu
13	×	[Glu, Ser, Tyr, Leu]	Ser.Leu.Tyr.Glu
14	×	[Glu, Tyr, Leu]	Tyr.Glu.Leu
15	× × × ×	[Glu, Gly, Val, Leu]	Gly.Ileu.Val.Glu
16	× × ×	Leu, Glu	Leu.Glu
17	× ×	[Glu, Val, Leu]	Ileu.Val.Glu
18	× ×	[Ser, Leu]	Ser.Leu
19	× ×	[Glu, Leu]	Pyr.Leu

* Leucine and isoleucine were not distinguished. Both are recorded as Leu.

Spot 1. No colour with ninhydrin. Located by slight fluorescence in ultraviolet light. It gave no DNP derivative, indicating no free —NH_2 group. *Spot* 2. R_F values and ionophoresis rates identical with those of peptide *Ap1*. *Spot* 3. Yellow with ninhydrin. R_F values and ionophoretic rates identical with those of peptide *Ap3*. *Spots* 4 *and* 6. Only detected when solvents run 2 lengths. May be deamidated forms of 3 and 5 or may have other structures. *Spot* 5. R_F values and ionophoretic rates identical with those of peptide *Ap6*. *Spot* 8. Gives traces of serine, valine and cysteic acid after hydrolysis. *Spots* 9, 13 *and* 18. Yellow with ninhydrin. *Spot* 10. Corresponds in composition to spots *Ap12* and *Am18* while spot 11 corresponds to spots *Ap11* and *Am17*, but the rates of movement in phenol and butanol–acetic acid were reversed. *Spot* 12. Obtained from the chromatogram was contaminated with spot 15, from which it could be separated by ionophoresis. Is probably identical with *Ap14*. *Spot* 14. Probably identical with *Ap15*. *Spot* 15. Strong grey colour with ninhydrin. *Spot* 19. No visible colour with ninhydrin. Located by fluorescence. Its R_F value in phenol is greater than the known peptides of leucine and glutamic acid. It is probably a pyrrolidone carboxylic acid derivative formed from an *N*-terminal glutamine residue.

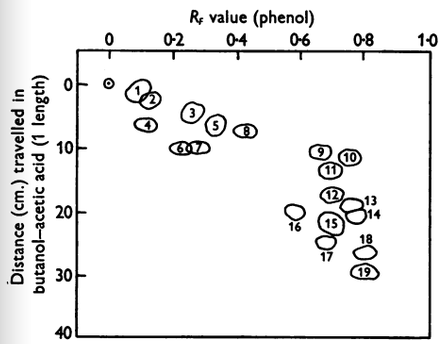


Fig. 3. Chromatogram of papain hydrolysate of fraction *A* (Expt. *Apa*) (see Table 2).

Table 3. *Partial hydrolysis products of peptide Apa1*

Peptide	Strength
$CySO_3H$.$CySO_3H$.Ala (*A*2δ1)	× ×
Cysteic acid	× ×
Glu.$CySO_3H$ (*A*1α1)	×
$CySO_3H$.Ala (*A*1γ3)	× ×
Glutamic acid	× ×
Ser.Val.$CySO_3H$ (*A*1ζ2)	× × × ×
Serine	×
Alanine	× × ×

carboxyl groups. The —SO_3H groups are considerably stronger and are probably completely ionized in this range. Fraction *A* (20·1 mg.) from which the bound acetate had been removed by washing well with 0·1N-HCl in acetone, was titrated to pH 2·27 with 0·025N-HCl. It was then titrated back to pH 7 with 0·025N-NaOH, taking the usual precautions. The results (Fig. 5) are corrected for the activity of HCl in water. Although there was considerable scatter at the lower pH values, it is evident that there are three free —COOH groups on fraction *A*. One of these is the free α-COOH group, so that two of the six dicarboxylic amino acids have free ω-COOH groups and the other four are present as amides.

Ionophoretic rates of peptides. Table 5 shows the distance moved by the various peptides on ionophoresis in 0·05M ammonium acetate. In this table the results listed in each column were obtained in the same experiment and so are directly comparable. Results from different experiments are not strictly comparable, although in general a peptide moving slightly towards the cathode is neutral and peptides moving towards the anode are acidic.

Table 6 shows the results of the amide determinations on the peptides produced by enzymic hydrolysis of fractions *A* and *B*.

Table 2. *Peptides from papain hydrolysate of fraction A* (*A*pa)

Spot no. (Fig. 3)	Approx. strength of peptide	Composition*	Probable structure
1	× × ×	[$CySO_3H$, Glu, Ser, Ala, Val]	Pyr.$CySO_3H$.$CySO_3H$.Ala.Ser.Val.$CySO_3H$
2	× ×	[$CySO_3H$, Asp, Glu, Tyr]	Glu.Asp.Tyr.$CySO_3H$.Asp
3	× ×	[$CySO_3H$, Asp, Tyr]	Asp.Tyr.$CySO_3H$.Asp
4	×	[$CySO_3H$, Asp, Tyr]	—
5	× × ×	[$CySO_3H$, Asp, Glu, Tyr, Leu]	Leu.Glu.Asp.Tyr.$CySO_3H$.Asp
6	×	[$CySO_3H$, Asp, Glu, Tyr, Leu]	—
7	×	—	Glutamic acid
8	× ×	—	Glycine
9	×	[Glu, Ser, Tyr]	?
10	× ×	—	Tyrosine
11	× × ×	[Glu, Tyr]	Tyr.Glu
12	×	[Glu, Tyr, Leu]	Tyr.Glu.Leu.Glu
13	×	[Glu, Ser, Tyr, Leu]	Ser.Leu.Tyr.Glu
14	×	[Glu, Tyr, Leu]	Tyr.Glu.Leu
15	× × × ×	[Glu, Gly, Val, Leu]	Gly.Ileu.Val.Glu
16	× × ×	Leu, Glu	Leu.Glu
17	× ×	[Glu, Val, Leu]	Ileu.Val.Glu
18	× ×	[Ser, Leu]	Ser.Leu
19	× ×	[Glu, Leu]	Pyr.Leu

* Leucine and isoleucine were not distinguished. Both are recorded as Leu.

Spot 1. No colour with ninhydrin. Located by slight fluorescence in ultraviolet light. It gave no DNP derivative, indicating no free —NH_2 group. *Spot* 2. R_F values and ionophoresis rates identical with those of peptide *A*p1. *Spot* 3. Yellow with ninhydrin. R_F values and ionophoretic rates identical with those of peptide *A*p3. *Spots* 4 *and* 6. Only detected when solvents run 2 lengths. May be deamidated forms of 3 and 5 or may have other structures. *Spot* 5. R_F values and ionophoretic rates identical with those of peptide *A*p6. *Spot* 8. Gives traces of serine, valine and cysteic acid after hydrolysis. *Spots* 9, 13 *and* 18. Yellow with ninhydrin. *Spot* 10. Corresponds in composition to spots *A*p12 and *A*m18 while spot 11 corresponds to spots *A*p11 and *A*m17, but the rates of movement in phenol and butanol–acetic acid were reversed. *Spot* 12. Obtained from the chromatogram was contaminated with spot 15, from which it could be separated by ionophoresis. Is probably identical with *A*p14. *Spot* 14. Probably identical with *A*p15. *Spot* 15. Strong grey colour with ninhydrin. *Spot* 19. No visible colour with ninhydrin. Located by fluorescence. Its R_F value in phenol is greater than the known peptides of leucine and glutamic acid. It is probably a pyrrolidone carboxylic acid derivative formed from an *N*-terminal glutamine residue.

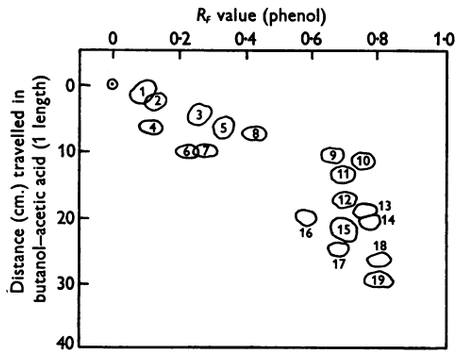


Fig. 3. Chromatogram of papain hydrolysate of fraction *A* (Expt. *A*pa) (see Table 2).

Table 3. *Partial hydrolysis products of peptide A*pa1

Peptide	Strength
$CySO_3H$.$CySO_3H$.Ala (*A*2δ1)	× ×
Cysteic acid	× ×
Glu.$CySO_3H$ (*A*1α1)	×
$CySO_3H$.Ala (*A*1γ3)	× ×
Glutamic acid	× ×
Ser.Val.$CySO_3H$ (*A*1ζ2)	× × × ×
Serine	×
Alanine	× × ×

carboxyl groups. The —SO_3H groups are considerably stronger and are probably completely ionized in this range. Fraction *A* (20·1 mg.) from which the bound acetate had been removed by washing well with 0·1 N-HCl in acetone, was titrated to pH 2·27 with 0·025 N-HCl. It was then titrated back to pH 7 with 0·025 N-NaOH, taking the usual precautions. The results (Fig. 5) are corrected for the activity of HCl in water. Although there was considerable scatter at the lower pH values, it is evident that there are three free —COOH groups on fraction *A*. One of these is the free α-COOH group, so that two of the six dicarboxylic amino acids have free ω-COOH groups and the other four are present as amides.

Ionophoretic rates of peptides. Table 5 shows the distance moved by the various peptides on ionophoresis in 0·05 M ammonium acetate. In this table the results listed in each column were obtained in the same experiment and so are directly comparable. Results from different experiments are not strictly comparable, although in general a peptide moving slightly towards the cathode is neutral and peptides moving towards the anode are acidic.

Table 6 shows the results of the amide determinations on the peptides produced by enzymic hydrolysis of fractions *A* and *B*.

Table 4. *Action of papain on peptide A*p5 (*A*p5pa)

Band no. (Fig. 4)	Strength	Composition	Probable structure
1	× × ×	[Glu, Gly, Val, Leu or Ileu]	Gly.Ileu.Val.Glu
2	×	[Tyr]	Free tyrosine
3	× × ×	[$CySO_3H$, Ser, Val, Leu or Ileu]	Val.$CySO_3H$.Ser.Leu
4	×	[$CySO_3H$, Glu, Ser, Gly, Ala, Val, Leu or Ileu]	Unchanged *A*p5
5	× × × ×	[$CySO_3H$, Glu, Ser, Ala]	Pyr.$CySO_3H$.$CySO_3H$.Ala.Ser

Band 3. Also contained traces of glutamic acid.

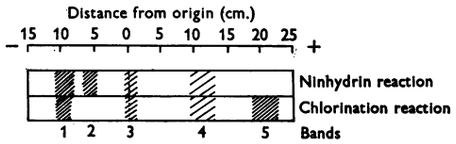


Fig. 4. Ionophoresis of papain hydrolysate of peptide *A*p5 in 0·2 M acetic acid, 220 v, 16 hr. (Expt. *A*p5pa) (see Table 4).

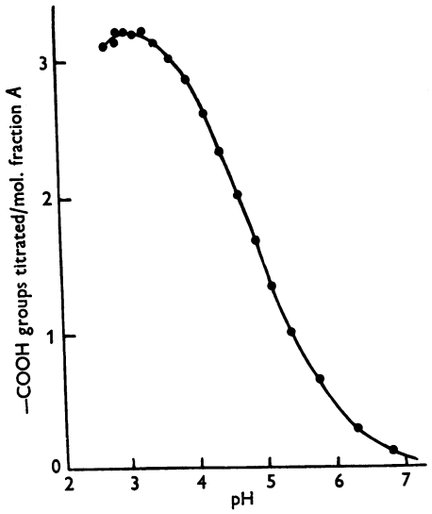


Fig. 5. Titration curve for fraction *A*.

DISCUSSION

Amide groups in the phenylalanyl chain

The amide contents of the peptides from fraction *B* (Table 6) were obtained using the fractions obtained by ionophoresis (Fig. 1), which may not have been completely pure. There may therefore be some error in the results but they are sufficiently accurate for the present purposes.

Sanger (1949*b*) showed that the aspartic and glutamic acid residues in positions 3 and 4 respectively were both in the form of amides, since several different *N*-terminal DNP-peptides containing the same amino acids could be obtained. This is supported by the amide content of peptide *B*p3, which gives a value approximating to the expected two residues.

Sanger & Tuppy (1951*b*) concluded that the glutamic acid residue in position 13 (see Table 7) contained a free γ-COOH group, since peptides *B*p1 (Val.Glu.Ala + Val.Glu) and *B*p5 (Leu.Val.Glu) from the peptic hydrolysates had the same R_F values as the corresponding peptides from the acid hydrolysates. This is confirmed by the fact that the peptides *B*p1 and *B*p2 (Val.Glu.Ala.Leu) are acidic (Fig. 1) and contain no amide groups (Table 6).

The remaining glutamic acid residue (position 21) is present in peptide *B*p4 (Leu.Val.$CySO_3H$.Gly.Glu.Arg.Gly.Phe). This is slightly acidic (Fig. 1) and contains no amide groups, so that the glutamic acid residue must have a free γ-COOH group.

Amide groups in the glycyl chains

It has already been shown that the *C*-terminal residue is asparagine (Harris, 1952; Sanger & Thompson, 1953*b*) and this is confirmed by the presence of one amide group in peptide *A*c1 ($CySO_3H$.Asp).

Several of the peptides obtained from the peptic hydrolysate have duplicate spots which contain the same amino acids, but are more acidic. Thus peptides *A*p10 (Tyr.Glu), *A*p4 (Asp.Tyr.$CySO_3H$.Asp), *A*p2 (Glu.Asp.Tyr.$CySO_3H$.Asp) and *A*p7 (Leu.Glu.Asp.Tyr.$CySO_3H$.Asp) are the corresponding forms of peptides *A*p11, *A*p3, *A*p1 and *A*p6 respectively. It was at first thought that this was due to deamidation occurring under the action of pepsin, but no loss of amide-N could be detected from fraction *A* when it was incubated with pepsin. More probably the insulin used was already partially deamidated, since Harfenist (1953) has shown that certain preparations contain a component having one less amide group than the normal insulin.

Peptides *A*p11 and *A*m17 both contain the sequence Tyr.Glu and are neutral. They are therefore Tyr.$GluNH_2$ and the glutamic acid residue in position 15 is present as an amide. There is considerable confirmation for this. Peptides *A*p15, *A*pa14 (Tyr.Glu.Leu) are also neutral and *A*p15 contains one amide group. Peptide *A*m15 (Glu.Leu.Glu) moves more slowly on ionophoresis than Leu.Glu (*A*m16) and must therefore be $GluNH_2$.Leu.Glu.

Table 5. *Ionophoretic rates of peptides in* 0·05 M *ammonium acetate*

The figures listed in each column represent the results obtained in a single experiment.

Peptide	Spot no.	Distance moved towards anode (cm.)								
		(15)*	(10)*	(15)*	(15)*	(14)*	(8·5)*	(15)*	(18)*	(18)*
Tyr.Glu	*Ap*11	—	−1	−2	—	—	—	—	—	—
	*Ap*10	—	—	5·5	—	—	—	—	—	—
	*Am*17	—	—	—	—	—	−1	—	—	—
Tyr.Glu.Leu	*Ap*15	—	−1	−2	—	—	—	—	—	—
	*Apa*14	—	—	—	—	—	—	−1	—	—
Tyr.Glu.Leu.Glu	*Ap*14	—	4·4	—	—	—	—	—	—	—
	*Apa*12	—	—	—	—	—	—	3·6	—	8
Leu.Glu	*Apa*16	—	—	—	—	—	—	6	—	—
	*Am*16	—	—	—	—	8·5	—	—	—	—
Glu.Leu.Glu	*Am*15	—	—	—	—	7	—	—	—	—
Asp.Tyr.$CySO_3H$.Asp	*Ap*3	8·7	7	—	8·2	—	—	—	—	—
	*Ap*4	15	—	14	—	—	—	—	—	—
	*Apa*3	—	—	—	8	—	—	—	—	—
Glu.Asp.Tyr.$CySO_3H$.Asp	*Ap*1	12·3	9	—	11·6	—	—	—	—	—
	*Ap*2	18·7	—	—	—	—	—	—	—	—
	*Apa*2	—	—	—	11·7	—	—	—	—	—
Leu.Glu.Asp.Tyr.-$CySO_3H$.Asp	*Ap*6	—	—	9	9·5	—	—	—	—	—
	*Ap*7	15	—	—	—	—	—	—	—	—
	*Apa*5	—	—	—	9·7	—	—	—	8	—
Gly.Ileu.Val.Glu	*Ap*16	—	—	4·3	—	—	—	—	—	—
	*Apa*15	—	—	—	—	—	—	4·3	—	—
	*Am*20	—	—	—	—	—	5	—	—	—
Ileu.Val.Glu	*Apa*17	—	—	—	6	—	—	—	—	—
Val.Glu	*Am*14	—	—	—	9	9·5	—	—	—	—
Ser.Leu.Tyr.Glu	*Apa*13	—	—	—	—	—	—	−1	—	—
Ser.Leu.Tyr	*Am*22	—	—	—	—	—	−1	—	—	—
Glu.Leu.Glu.Asp.Tyr.-$CySO_3H$.Asp	*Ac*3	—	—	—	—	—	—	—	6·7	—
Glu.Leu.Glu.Asp.Tyr	*Ac*5	—	—	—	—	—	—	—	—	6·7
Glutamine	*Am*10	—	—	—	—	−1	—	—	—	—
Asparagine	*Am*9	—	—	—	—	−1	—	—	—	—
Glutamic acid	—	14·5	—	16	—	—	—	—	—	—

* Time of ionophoresis (hr.).

Table 6. *Amide determinations of peptides*

Peptide	Sequence	No. of estimations averaged	Amide groups (residues/mol.)
*B*p1	Val.Glu.Ala	2	0·15
*B*p2	Val.Glu.Ala.Leu	2	0·15
*B*p4	Leu.Val.$CySO_3H$.Gly.Glu.Arg.Gly.Phe	3	0·13
*B*p3	Phe.Val.Asp.Glu.His.Leu.$CySO_3H$.Gly.Ser.His.Leu	2	1·7
*A*c1	$CySO_3H$.Asp	1	0·7
*A*p15	Tyr.Glu.Leu	2	0·98
*A*p14	Tyr.Glu.Leu.Glu	2	1·03
*A*p3	Asp.Tyr.$CySO_3H$.Asp	4	2·10
*A*p1	Glu.Asp.Tyr.$CySO_3H$.Asp	7	1·94
*A*p5pa1	Gly.Ileu.Val.Glu	2	0·15
*A*p5	Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala.Ser.Val.-$CySO_3H$.Ser.Leu	17	1·16

Similarly, Glu.Leu.Glu.Asp.Tyr.$CySO_3H$.Asp (*A*c3) moves slower than Leu.Glu.Asp.Tyr.-$CySO_3H$.Asp (*A*pa5).

In the chymotryptic digests two peptides (*A*c5, *A*c7) were isolated having the same amino acid composition but one (*A*c7) gave no colour with ninhydrin. It was pointed out that this was presumably due to the lability of an *N*-terminal glutaminyl residue which tends to cyclize to a pyrrolidone carboxylic acid residue. Thus *A*c5 is probably $GluNH_2$.Leu.Glu.Asp.Tyr and *A*c7 Pyr.Leu.Glu.Asp.Tyr. Also Pyr.Leu (*A*pa19) was found in the papain hydrolysate.

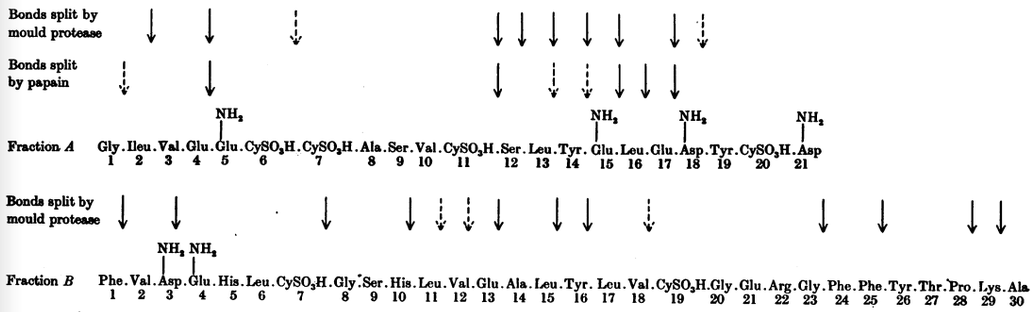
Whereas peptide *A*p15 (Tyr.$GluNH_2$.Leu) was found to be neutral, peptide *A*p14 (Tyr.Glu.Leu.-Glu) is acidic, which suggests the second glutamic acid residue has a free carboxyl group. This is supported by the fact that Leu.Glu (*A*pa16, *A*m16) and Glu.Leu.Glu (*A*m15) are acidic. The glutamic acid residue in position 17 therefore contains a free γ-COOH group. Amide determination also showed that peptide *A*p14 (Tyr.$GluNH_2$.Leu.Glu) had only one amide group, as did *A*p15 (Tyr.Glu-NH_2.Leu). Peptides *A*p1 and *A*pa2 (Glu.Asp.Tyr.-$CySO_3H$.Asp) differ from peptides *A*p3 and *A*pa3 (Asp.Tyr.$CySO_3H$.Asp) only by one glutamic acid residue but are more acidic on ionophoresis, which confirms that this glutamic acid residue (position 17) is not amidized. Similarly peptide *A*p6 (Leu.Glu.-Asp.Tyr.$CySO_3H$.Asp) is less acidic than *A*p1 and more acidic than *A*p3.

Peptide *A*p3 (Asp.Tyr.$CySO_3H$.Asp) contains two amide groups (Table 6). These can only be located on the aspartic acid residues so that both (positions 18 and 21) must be present as asparagine. Similarly *A*p1 (Glu.Asp.Tyr.$CySO_3H$.Asp) contains two amide residues, and the glutamic acid residue has a free γ-COOH group. Further evidence for an amide in position 18 is that Glu.Leu.Glu.-Asp.Tyr (*A*c5) moves slower on ionophoresis than Tyr.Glu.Leu.Glu (*A*p14). These two peptides differ only by the presence of the aspartic acid residue in question, although they contain a different tyrosine residue.

Three enzymes have split the bond between the two glutamic acid residues in positions 4 and 5. Thus, papain gave a high yield of peptide *A*pa15 (Gly.Ileu.Val.Glu). On ionophoresis it moved towards the anode, showing that the glutamic acid was not present as an amide. Similarly, *A*p16 and *A*m20, which have the same sequence as *A*pa15 are acidic and NH_3 determination showed there was no amide in *A*p5pa1. Peptides *A*pa17 (Ileu.Val.Glu) and *A*m14 (Val.Glu) are also acidic. Peptide *A*p5pa5 was identified as a pyrrolidonoyl peptide (Pyr.$CySO_3H$.$CySO_3H$.Ala.Ser) indicating that the residue in position 5 is glutamine.

From these results it was expected that peptide *A*p5 (Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.-Ala.Ser.Val.$CySO_3H$.Ser.Leu) would contain one amide group (on position 5). However, innumerable determinations gave values which varied somewhat but were considerably closer to two than to one. Also a few determinations on peptides *A*p7, *A*c2 and *A*c4 gave similar high results. The *A*p5 was prepared either from chromatograms or more often by a single ionophoresis in 0·2 M acetic acid. Although the apparent absence of aspartic acid and tyrosine from hydrolysates suggested that the *A*p5 was reasonably pure, it now seems that the high amide value was due to the presence of an amide-rich impurity, since when *A*p5 was further purified by high-voltage ionophoresis at pH 3·7, values approximating to one were obtained. The figures averaged in Table 6 are all for such purified material. This impurity appears to travel slightly slower than *A*p5 in the 0·2 M acetic acid ionophoresis since a cut taken behind the *A*p5 band gave a higher amide value than the main band. It seems probable that it is largely unchanged fraction *A* and possibly

Table 7. *Amide distribution in insulin*



other similar large peptides. Fraction *A* has a comparatively high amide content and a small amount present in the *A*p5 would cause considerable error. In our experience it forms very long spread-out streaks both on ionophoresis and chromatography and, if present, would be expected to contaminate *A*p5 and other peptides (*A*p9, *A*c2, *A*c4) which move at similar rates.

On the basis of these combined results it is concluded that the distribution of amide groups in insulin is as shown in Table 7. This agrees well with the results of Chibnall & Rees (1952), who showed that all three aspartic acid and three of the glutamic acid residues are amidized.

The possibility of iso*glutamine residues*

Kandel, Kandel, Kovacs & Bruckner (1954) have recently suggested that insulin contains an *iso*glutamine residue. This is based on the finding that when insulin is subjected to the Hofmann degradation and hydrolysed, a substance is produced which gives a 2:4-dinitrophenylhydrazone that moves at the same rate on a butanol–3 % NH_3 paper chromatogram as the dinitrophenylhydrazone of 4-oxobutyric acid. This product would be obtained if an *iso*glutaminyl residue were present in the protein. We have repeated the above experiment and have confirmed the presence of a spot that moves on a butanol–3 % NH_3 chromatogram at the same rate as the synthetic dinitrophenylhydrazone of 4-oxobutyric acid. There were, however, rather more spots on the chromatogram than the two recorded by Kandel *et al.* Samples of the dinitrophenylhydrazone spot were eluted from the paper and reinvestigated by chromatography on *tert.*-amyl alcohol–phthalate (pH 6) buffer (Blackburn & Lowther, 1951) and in the 'toluene' solvent of Biserte & Osteux (1951). With the former solvent two strong spots were produced which had moved 6 and 11 cm. down the paper respectively and a fainter spot that had moved 19 cm. A control spot of the synthetic hydrazone had moved 19 cm. Similar results were obtained with the 'toluene' solvent. It thus seems that the spot that was identified by Kandel *et al.* as the above hydrazone, was in fact a mixture of several substances, one of which may have been the hydrazone. It is present in only small quantities and would probably account for about 1–2 % of one residue in insulin (mol.wt. 5700). In view of these results and the known complexity of the reactions of hypochlorite with amino acids, peptides and proteins (see e.g. Goldschmidt & Strauss, 1930 and previous papers) it seems that there is no clear evidence for the existence of *iso*glutamine residues in insulin.

The fact that the residues in positions *A*5 and *A*15 are converted to a pyrrolidonoyl group when they are the *N*-terminal residues of a peptide (e.g. in peptides *A*pa1 and *A*c7, respectively) indicates that they are glutaminyl residues. An *iso*glutaminyl residue would not decompose to a pyrrolidonoyl peptide but would give the amide of pyrrolidone carboxylic acid. No such evidence exists for the residue in position *B*4, but a study of the Hofmann degradation of fractions *A* and *B* of oxidized insulin showed that the spot moving at the same rate as the dinitrophenylhydrazone of 4-oxobutyric acid was produced from both fractions.

Action of the proteolytic enzymes

The main sites at which splitting occurred under the action of papain and mould protease are indicated in Table 7. The sites of action of mould protease on fraction *B* are based on results of a single experiment and should not be regarded as conclusive and are not comprehensive. Both enzymes show a wide specificity which is somewhat similar to that of pepsin, though the relative intensities of splitting of the various bonds was markedly different for the three enzymes. Mould protease carried the digestion of fraction *A* further than the other enzymes. For structural studies in the future, this enzyme may prove useful, for example, in degrading longer chains obtained by the use of other enzymes.

It is doubtful if the preparations of the two proteinases represent pure enzymes. In the case of papain we have also studied the action of an activated sample of crystalline mercuripapain prepared by Kimmel & Smith (1954). The specificity appeared to be essentially the same as that found with the crude commercial sample.

It is interesting that papain breaks peptide *A*p5 in a different way from fraction *A*. In fraction *A* strong splitting occurs at the $CySO_3H$.Ser bond (position 11–12), whereas this bond was stable in *A*p5 and splitting occurred largely at the Ser.Val bond (position 9–10). This result could be explained if papain was inhibited by a free carboxyl group on the second residue from the bond. Other explanations are possible, but it does illustrate that the specificity of proteolytic enzymes is not conditioned solely by the nature of the residues involved in the bond in question.

SUMMARY

1. Fraction *A* of oxidized insulin was hydrolysed with mould protease and cyanide-activated papain and the resulting peptides studied. Both enzymes showed a rather wide specificity.

2. The positions of the amide groups of insulin were determined by estimating the relative ionophoretic mobilities and the amide contents of peptides from enzymic hydrolysates of fractions *A* and *B* of oxidized insulin. The results are summarized in Table 7.

We wish to thank Dr R. R. Porter for the mould protease and Dr Emil L. Smith for the crystalline mercuripapain.

One of us (E. O. P. T.) wishes to thank the Commonwealth Government of Australia for a Commonwealth Scientific and Industrial Research Organization studentship.

REFERENCES

Bailey, K. (1937). *Biochem. J.* **31**, 1406.
Biserte, G. & Osteux, R. (1951). *Bull. Soc. Chim. biol., Paris*, **33**, 50.
Blackburn, S. & Lowther, A. G. (1951). *Biochem. J.* **48**, 126.
Chibnall, A. C. & Rees, M. W. (1952). *Biochem. J.* **52**, iii.
Conway, E. J. (1950). *Microdiffusion Analysis and Volumetric Error.* London: Lockwood.
Crewther, W. G. & Lennox, F. G. (1950). *Nature, Lond.*, **165**, 680.
Durrum, E. L. (1950). *J. Amer. chem. Soc.* **72**, 2943.
Goldschmidt, S. & Strauss, K. (1930). *Ber. dtsch. chem. Ges.* **63**, 1218.
Harfenist, E. J. (1953). *J. Amer. chem. Soc.* **75**, 5528.
Harris, J. I. (1952). *J. Amer. chem. Soc.* **74**, 2944.
Kandel, I., Kandel, M., Kovacs, J. & Bruckner, V. (1954). *Naturwissenschaften*, **41**, 281.
Kimmel, J. R. & Smith, E. L. (1954). *J. biol. Chem.* **207**, 515.
Michl, H. (1951). *Monatshefte*, **82**, 489.
Moore, S. & Stein, W. H. (1948). *J. biol. Chem.* **176**, 367.
Moore, S. & Stein, W. H. (1951). *J. biol. Chem.* **192**, 663.
Rees, M. W. (1946). *Biochem. J.* **40**, 632.
Rydon, H. N. & Smith, P. W. G. (1952). *Nature, Lond.*, **169**, 922.
Sanger, F. (1949*a*). *Biochem. J.* **44**, 126.
Sanger, F. (1949*b*). *Biochem. J.* **45**, 563.
Sanger, F. & Thompson, E. O. P. (1953*a*). *Biochem. J.* **53**, 353.
Sanger, F. & Thompson, E. O. P. (1953*b*). *Biochem. J.* **53**, 366.
Sanger, F. & Tuppy, H. (1951*a*). *Biochem. J.* **49**, 463.
Sanger, F. & Tuppy, H. (1951*b*). *Biochem. J.* **49**, 481.

Metabolism of Polycyclic Compounds

8.* ACID-LABILE PRECURSORS OF NAPHTHALENE PRODUCED AS METABOLITES OF NAPHTHALENE

By E. BOYLAND and J. B. SOLOMON
Chester Beatty Research Institute, Institute of Cancer Research, Royal Cancer Hospital, Fulham Road, London, S.W. 3

(*Received* 9 *October* 1954)

Baumann & Herter (1877) observed that crystals of naphthalene appeared in the distillate when the acidified urine of rabbits dosed with naphthalene was heated. Bourne & Young (1933) noticed that crystals of naphthalene separated from such acidified urine in the cold. They suggested that this might be due to the dehydration of 1:2-dihydro-2-naphthol, which is known to give naphthalene with cold mineral acid (Bamberger & Lodter, 1895), but dihydronaphthol does not appear to have been isolated from urines of animals dosed with naphthalene. Chang & Young (1943) and Young (1947) estimated that the acid-liberated naphthalene accounted for from 4 to 14 % of the dose fed to rats or rabbits. The liberation of anthracene on acidification of the urine of rats and rabbits dosed with anthracene was noticed by Boyland & Levi (1936). Chang & Young (1943) also found that small amounts of hydrocarbons were liberated on acidification of urines of rats dosed with phenanthrene and anthracene. Acid-labile precursors could not be detected in the urines of rats dosed with higher polycyclic hydrocarbons.

* No. 7 of this series: Boyland, E. & Wiltshire, G. M. (1953).

Studies on 1:2:3:4-tetrahydronaphthalene (tetralin) metabolism by Pohl & Rawicz (1919) and by Rockemann (1922) indicated that tetralin was converted into 1- or 2-tetralyl glucosiduronic acids by dogs and rabbits. When these tetralyl glucosiduronic acids were steam-distilled from acid solution, 1:2-dihydronaphthalene and naphthalene were produced. Pohl & Rawicz (1919) also dosed rabbits with 1:2-dihydronaphthalene and found that naphthalene was produced on acidification of the urine; they suggested that a 1:2-dihydronaphthyl glucosiduronic acid had been formed. Corner, Billett & Young (1954) have recently shown that 1:2-dihydronaphthalene-1:2-diol is excreted as a glucosiduronic acid derivative by rabbits dosed with naphthalene.

Evidence that naphthalene is metabolized to 1:2-dihydro-1-naphthyl glucosiduronic acids, which readily yield naphthalene in acid solution, is presented in this paper.

EXPERIMENTAL

Fifty rats (average body weight 250 g.) maintained on rat cake and water, were each injected intraperitoneally with naphthalene (100 mg.) dissolved in arachis oil (1 ml.) on

azone spot were eluted from the
nvestigated by chromatography on
nol–phthalate (pH 6) buffer (Black-
r, 1951) and in the 'toluene' solvent
teux (1951). With the former solvent
ts were produced which had moved 6
own the paper respectively and a
at had moved 19 cm. A control spot
tic hydrazone had moved 19 cm.
were obtained with the 'toluene'
us seems that the spot that was
andel *et al.* as the above hydrazone,
nixture of several substances, one of
e been the hydrazone. It is present
quantities and would probably
out 1–2% of one residue in insulin
. In view of these results and the
xity of the reactions of hypochlorite
ids, peptides and proteins (see e.g.
Strauss, 1930 and previous papers)
there is no clear evidence for the
glutamine residues in insulin.
the residues in positions *A*5 and *A*15
o a pyrrolidonoyl group when they are
residues of a peptide (e.g. in peptides
the crude commercial sample.

It is interesting that papain breaks peptide *A*p5 in a different way from fraction *A*. In fraction *A* strong splitting occurs at the $CySO_3H$.Ser bond (position 11–12), whereas this bond was stable in *A*p5 and splitting occurred largely at the Ser.Val bond (position 9–10). This result could be explained if papain was inhibited by a free carboxyl group on the second residue from the bond. Other explanations are possible, but it does illustrate that the specificity of proteolytic enzymes is not conditioned solely by the nature of the residues involved in the bond in question.

SUMMARY

1. Fraction *A* of oxidized insulin was hydrolysed with mould protease and cyanide-activated papain and the resulting peptides studied. Both enzymes showed a rather wide specificity.

2. The positions of the amide groups of insulin were determined by estimating the relative iono-phoretic mobilities and the amide contents of peptides from enzymic hydrolysates of fractions *A* and *B* of oxidized insulin. The results are summarized in Table 7.

Table 5. *Ionophoretic rates of peptides in* 0·05 M *ammonium acetate*

The figures listed in each column represent the results obtained in a single experiment.

Peptide	Spot no.	Distance moved towards anode (cm.)								
		(15)*	(10)*	(15)*	(15)*	(14)*	(8·5)*	(15)*	(18)*	(18)*
Tyr.Glu	*Ap*11	—	−1	−2	—	—	—	—	—	—
	*Ap*10	—	—	5·5	—	—	—	—	—	—
	*Am*17	—	—	—	—	—	−1	—	—	—
Tyr.Glu.Leu	*Ap*15	—	−1	−2	—	—	—	—	—	—
	*Apa*14	—	—	—	—	—	—	−1	—	—
Tyr.Glu.Leu.Glu	*Ap*14	—	4·4	—	—	—	—	—	—	—
	*Apa*12	—	—	—	—	—	—	3·6	—	8
Leu.Glu	*Apa*16	—	—	—	—	—	—	6	—	—
	*Am*16	—	—	—	—	8·5	—	—	—	—
Glu.Leu.Glu	*Am*15	—	—	—	—	7	—	—	—	—
Asp.Tyr.$CySO_3H$.Asp	*Ap*3	8·7	7	—	8·2	—	—	—	—	—
	*Ap*4	15	—	14	—	—	—	—	—	—
	*Apa*3	—	—	—	8	—	—	—	—	—
Glu.Asp.Tyr.$CySO_3H$.Asp	*Ap*1	12·3	9	—	11·6	—	—	—	—	—
	*Ap*2	18·7	—	—	—	—	—	—	—	—
	*Apa*2	—	—	—	11·7	—	—	—	—	—
Leu.Glu.Asp.Tyr.-$CySO_3H$.Asp	*Ap*6	—	—	9	9·5	—	—	—	—	—
	*Ap*7	15	—	—	—	—	—	—	—	—
	*Apa*5	—	—	—	9·7	—	—	—	8	—
Gly.Ileu.Val.Glu	*Ap*16	—	—	4·3	—	—	—	—	—	—
	*Apa*15	—	—	—	—	—	—	4·3	—	—
	*Am*20	—	—	—	—	—	5	—	—	—
Ileu.Val.Glu	*Apa*17	—	—	—	6	—	—	—	—	—
Val.Glu	*Am*14	—	—	—	9	9·5	—	—	—	—
Ser.Leu.Tyr.Glu	*Apa*13	—	—	—	—	—	—	−1	—	—
Ser.Leu.Tyr	*Am*22	—	—	—	—	—	−1	—	—	—
Glu.Leu.Glu.Asp.Tyr.-$CySO_3H$.Asp	*Ac*3	—	—	—	—	—	—	—	6·7	—
Glu.Leu.Glu.Asp.Tyr	*Ac*5	—	—	—	—	—	—	—	—	6·7
Glutamine	*Am*10	—	—	—	—	−1	—	—	—	—
Asparagine	*Am*9	—	—	—	—	−1	—	—	—	—
Glutamic acid	—	14·5	—	16	—	—	—	—	—	—

* Time of ionophoresis (hr.).

Table 6. *Amide determinations of peptides*

Peptide	Sequence	No. of estimations averaged	Amide groups (residues/mol.)
*Bp*1	Val.Glu.Ala	2	0·15
*Bp*2	Val.Glu.Ala.Leu	2	0·15
*Bp*4	Leu.Val.$CySO_3H$.Gly.Glu.Arg.Gly.Phe	3	0·13
*Bp*3	Phe.Val.Asp.Glu.His.Leu.$CySO_3H$.Gly.Ser.His.Leu	2	1·7
*Ac*1	$CySO_3H$.Asp	1	0·7
*Ap*15	Tyr.Glu.Leu	2	0·98
*Ap*14	Tyr.Glu.Leu.Glu	2	1·03
*Ap*3	Asp.Tyr.$CySO_3H$.Asp	4	2·10
*Ap*1	Glu.Asp.Tyr.$CySO_3H$.Asp	7	1·94
*Ap*5pa1	Gly.Ileu.Val.Glu	2	0·15
*Ap*5	Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.Ala.Ser.Val.-$CySO_3H$.Ser.Leu	17	1·16

Similarly, Glu.Leu.Glu.Asp.Tyr.$CySO_3H$.Asp (*Ac*3) moves slower than Leu.Glu.Asp.Tyr.-$CySO_3H$.Asp (*Apa*5).

In the chymotryptic digests two peptides (*Ac*5, *Ac*7) were isolated having the same amino acid composition but one (*Ac*7) gave no colour with ninhydrin. It was pointed out that this was presumably due to the lability of an *N*-terminal glutaminyl residue which tends to cyclize to a pyrrolidone carboxylic acid residue. Thus *Ac*5 is probably $GluNH_2$.Leu.Glu.Asp.Tyr and *Ac*7 Pyr.Leu.Glu.Asp.Tyr. Also Pyr.Leu (*Apa*19) was found in the papain hydrolysate.

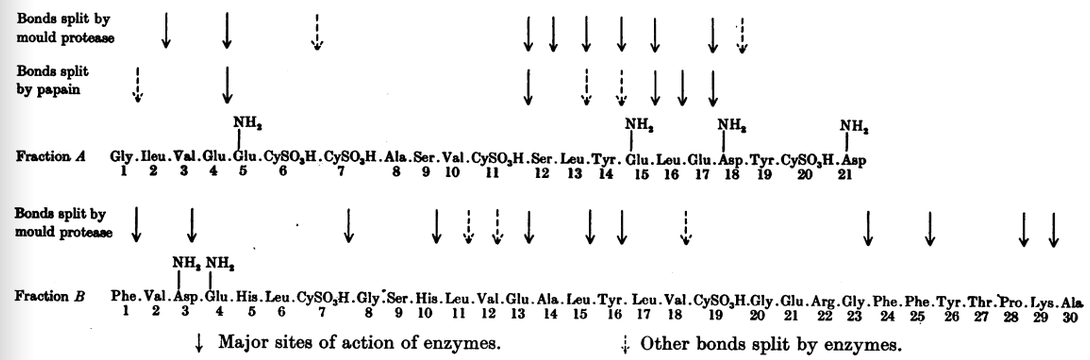
Whereas peptide *Ap*15 (Tyr.$GluNH_2$.Leu) was found to be neutral, peptide *Ap*14 (Tyr.Glu.Leu.-Glu) is acidic, which suggests the second glutamic acid residue has a free carboxyl group. This is supported by the fact that Leu.Glu (*Apa*16, *Am*16) and Glu.Leu.Glu (*Am*15) are acidic. The glutamic acid residue in position 17 therefore contains a free γ-COOH group. Amide determination also showed that peptide *Ap*14 (Tyr.$GluNH_2$.Leu.Glu) had only one amide group, as did *Ap*15 (Tyr.Glu-NH_2.Leu). Peptides *Ap*1 and *Apa*2 (Glu.Asp.Tyr.-$CySO_3H$.Asp) differ from peptides *Ap*3 and *Apa*3 (Asp.Tyr.$CySO_3H$.Asp) only by one glutamic acid residue but are more acidic on ionophoresis, which confirms that this glutamic acid residue (position 17) is not amidized. Similarly peptide *Ap*6 (Leu.Glu.-Asp.Tyr.$CySO_3H$.Asp) is less acidic than *Ap*1 and more acidic than *Ap*3.

Peptide *Ap*3 (Asp.Tyr.$CySO_3H$.Asp) contains two amide groups (Table 6). These can only be located on the aspartic acid residues so that both (positions 18 and 21) must be present as asparagine. Similarly *Ap*1 (Glu.Asp.Tyr.$CySO_3H$.Asp) contains two amide residues, and the glutamic acid residue has a free γ-COOH group. Further evidence for an amide in position 18 is that Glu.Leu.Glu.-Asp.Tyr (*Ac*5) moves slower on ionophoresis than Tyr.Glu.Leu.Glu (*Ap*14). These two peptides differ only by the presence of the aspartic acid residue in question, although they contain a different tyrosine residue.

Three enzymes have split the bond between the two glutamic acid residues in positions 4 and 5. Thus, papain gave a high yield of peptide *Apa*15 (Gly.Ileu.Val.Glu). On ionophoresis it moved towards the anode, showing that the glutamic acid was not present as an amide. Similarly, *Ap*16 and *Am*20, which have the same sequence as *Apa*15 are acidic and NH_3 determination showed there was no amide in *Ap*5pa1. Peptides *Apa*17 (Ileu.Val.Glu) and *Am*14 (Val.Glu) are also acidic. Peptide *Ap*5pa5 was identified as a pyrrolidonoyl peptide (Pyr.$CySO_3H$.$CySO_3H$.Ala.Ser) indicating that the residue in position 5 is glutamine.

From these results it was expected that peptide *Ap*5 (Gly.Ileu.Val.Glu.Glu.$CySO_3H$.$CySO_3H$.-Ala.Ser.Val.$CySO_3H$.Ser.Leu) would contain one amide group (on position 5). However, innumerable determinations gave values which varied somewhat but were considerably closer to two than to one. Also a few determinations on peptides *Ap*7, *Ac*2 and *Ac*4 gave similar high results. The *Ap*5 was prepared either from chromatograms or more often by a single ionophoresis in 0·2 M acetic acid. Although the apparent absence of aspartic acid and tyrosine from hydrolysates suggested that the *Ap*5 was reasonably pure, it now seems that the high amide value was due to the presence of an amide-rich impurity, since when *Ap*5 was further purified by high-voltage ionophoresis at pH 3·7, values approximating to one were obtained. The figures averaged in Table 6 are all for such purified material. This impurity appears to travel slightly slower than *Ap*5 in the 0·2 M acetic acid ionophoresis since a cut taken behind the *Ap*5 band gave a higher amide value than the main band. It seems probable that it is largely unchanged fraction *A* and possibly

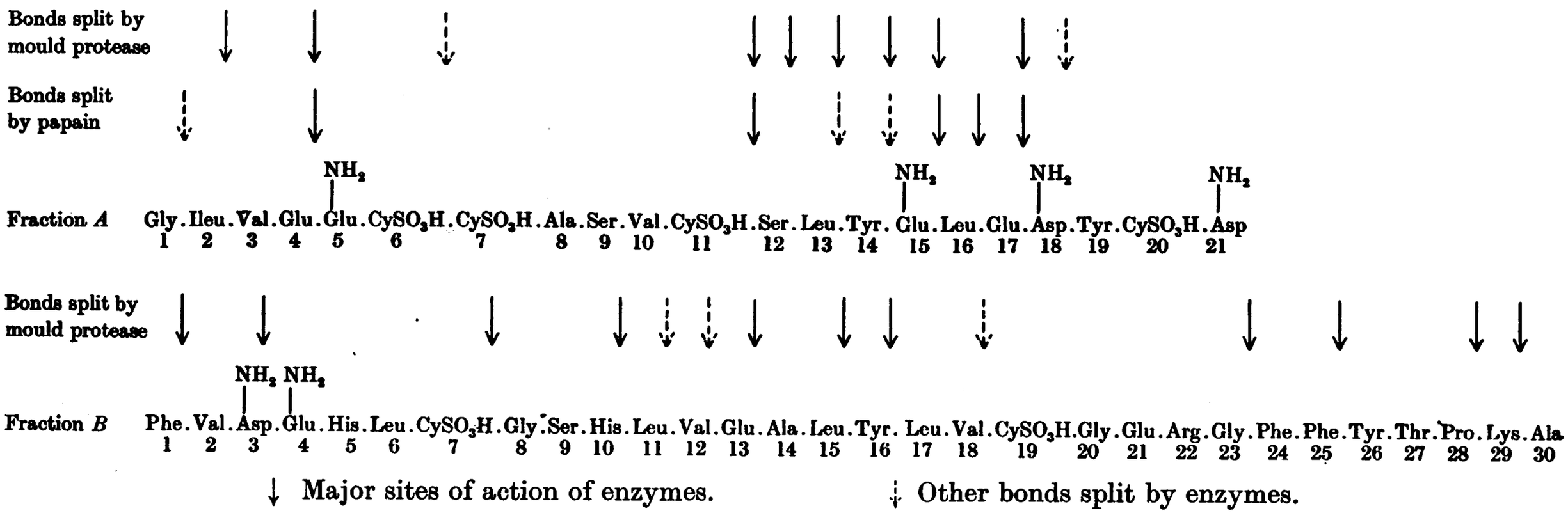
Table 7. *Amide distribution in insulin*



is not amidized. Similarly peptide *A*p6 (Leu.Glu.Asp.Tyr.$CySO_3H$.Asp) is less acidic than *A*p1 and more acidic than *A*p3.

Peptide *A*p3 (Asp.Tyr.$CySO_3H$.Asp) contains two amide groups (Table 6). These can only be located on the aspartic acid residues so that both (positions 18 and 21) must be present as asparagine. Similarly *A*p1 (Glu.Asp.Tyr.$CySO_3H$.Asp) contains two amide residues, and the glutamic acid residue has a free γ-COOH group. Further evidence for an amide in position 18 is that Glu.Leu.Glu.Asp.Tyr (*A*c5) moves slower on ionophoresis than from hydrolysates suggested that the *A*p5 was reasonably pure, it now seems that the high amide value was due to the presence of an amide-rich impurity, since when *A*p5 was further purified by high-voltage ionophoresis at pH 3·7, values approximating to one were obtained. The figures averaged in Table 6 are all for such purified material. This impurity appears to travel slightly slower than *A*p5 in the 0·2 M acetic acid ionophoresis since a cut taken behind the *A*p5 band gave a higher amide value than the main band. It seems probable that it is largely unchanged fraction *A* and possibly

Table 7. *Amide distribution in insulin*



Eur. J. Biochem. **269**, 3894–3904 (2002) © FEBS 2002 doi:10.1046/j.1432-1033.2002.03055.x

Schemes of flux control in a model of *Saccharomyces cerevisiae* glycolysis

Leighton Pritchard and Douglas B. Kell

Institute of Biological Sciences, University of Wales, Aberystwyth, UK

We used parameter scanning to emulate changes to the limiting rate for steps in a fitted model of glucose-derepressed yeast glycolysis. Three flux-control regimes were observed, two of which were under the dominant control of hexose transport, in accordance with various experimental studies and other model predictions. A third control regime in which phosphofructokinase exerted dominant glycolytic flux control was also found, but it appeared to be physiologically unreachable by this model, and all realistically obtainable flux control regimes featured hexose transport as a step involving high flux control.

Keywords: yeast; metabolic control analysis; glycolysis; modelling; flux.

In vivo and *in vitro* investigations of metabolic pathways can be complex and expensive. The need to focus efficiently both monetary and physical effort necessitates that some pathways and organisms will be only partially explored by experiment, while others will be neglected completely. Bioinformatic and computational approaches offer a means of obtaining full value from experimentally acquired data, extending their interpretation, suggesting novel hypotheses for future experiments and guiding the experimentalist towards potentially rewarding investigations but away from likely fruitless ones. In this paper, we use such an approach, parameter scanning, to investigate the operation of a model of glucose derepressed yeast glycolysis (fitted by evolutionary computing to experimental data) under a far wider range of conditions than could be considered *in vitro* or *in vivo*, which suggests opportunities for further experiment.

Glycolysis is perhaps the most important pathway in the metabolism of many living cells, describing the conversion of glucose (and sometimes other hexoses) to pyruvate and thence, in some organisms, to ethanol. In this conversion two molecules of ATP are consumed and four molecules of ATP are generated, providing a major source of 'negotiable energy' for the cell. The glycolytic pathway, though crucial to each, varies in detail between organisms [1]; for largely economic reasons, greater effort has gone into the understanding of glycolysis in some organisms than in others. Brewer's yeast *Saccharomyces cerevisiae*, and in particular its glycolytic pathway, is of great economic importance, not least for the production of ethanol in the brewing and distilling industries. The study of yeast glycolysis has thus been the focus of scientific interest for over a century. In pregenomic studies, the enzymes and metabolites that make up the pathway were considered to have been elucidated completely [2], but the solution of the *S. cerevisiae* genome added further to this knowledge, and it is widely considered that this organism currently possesses the best-investigated and best-understood glycolytic pathway.

Much effort has already been invested in mathematical modelling of the glycolytic pathway in yeast [3–8] and in other organisms, such as *Trypanosoma brucei*, the parasite that causes sleeping sickness [9–12]. The success and utility of modelling in the study of *T. brucei* glycolysis has even led to the coining of a new strategy for the investigation of metabolism: computer experimentation [11]. This is intended to be a substitute for practical experimentation, and must be based on precise kinetic knowledge of the system. For yeast glycolysis, the most complete model to date was constructed in order to test whether combining the properties of the individual enzymes in isolation would yield a proper description of the pathway as a whole [7]. This work provided a unique and highly valuable set of *in vitro* kinetic and physical data obtained under a consistent set of conditions (rare in the field [13]), and represented a major step towards such computer experimentation in yeast.

In this paper, and in the spirit of computer experimentation, we use the parameter scanning functions of GEPASI [14–16] to generate over 8000 models of glucose-derepressed yeast glycolysis in order to test the flux-control characteristics of the Teusink *et al.* model [7] under a wide range of enzyme limiting rates. The limiting rates for 13 steps of the model were scanned independently in all combinations by an overall factor of four. In this way we explore the flux-control behaviour of the model within the limits of its description of glucose-derepressed glycolysis and fixed fluxes to glycogen and trehalose.

Although the *in vitro* kinetic data from [7] are rather precise and quite complete, the generated model was unable perfectly to describe the system's *in vivo* behaviour. The authors, however, were not aiming to give the best possible description of the experimental system, but were instead investigating whether the isolated, *in vitro* kinetics of the glycolytic enzymes could describe the experimental system. Nonetheless they attempted to fit individual steps to experimental data, but restrained themselves from attempting to fit simultaneously the whole model, and from presuming that the intracellular concentration of enzyme was the single cause of the discrepancy between *in vitro* and *in vivo* behaviour for each individual step; they also considered the effects of altered substrate/product affinities and equilibrium constants. It was seen that, for most of the enzymes, only a small change in the value of the limiting rate was required for *in silico* kinetics to match each individual enzyme's *in vivo* performance closely. While modifications of V_{max} alone to fit *in vivo* performance could be calculated analytically for most steps in glycolysis, this was not the case for all steps [7].

In this paper, we use a version of the model of glucose-derepressed wild type yeast glycolysis described in [7] and investigate characteristics of its operation close to the wild-type state, and over a much wider range of operation than that for which the model was originally intended. It has been suggested [17] that inductive, multivariate and machine learning approaches are appropriate for such problems, and so we used the evolutionary programming algorithms incorporated in the metabolic modelling package GEPASI [14–16] to estimate multiple sets of V_{max} values for the glycolytic enzymes that enable the model to describe *in vivo* behaviour closely. Such an approach, although unlike algebraic analysis in that it produces a range of possible (although inexact) fits to the data, accounts for the effect of simultaneously varying the kinetics of the other steps, and is also expected to be a better qualitative measure of the flexibility of the model itself in describing the experimental data than is algebraically fitting isolated steps to their *in vivo* performance. As a population operating approximately equally close to the observed experimental state in [7], these models may be considered to represent natural variability in the yeast population, and we investigated the regions of parameter and variable space described by them. Metabolic control analysis [18–21] was performed on the fitted models, and rank correlation analysis [22,23] used to investigate patterns of flux control. The model with the best-fit V_{max} parameters was used as the base model for parameter scanning using routines contained in GEPASI.

METHODS

Model

A model of branched glycolysis, as described in [7] was obtained in SCAMP format from one of its authors (a kind gift from B. Teusink, TNO Prevention and Health, Leiden, the Netherlands.) and is illustrated schematically in Fig. 1. The ordinary differential equations describing the model are

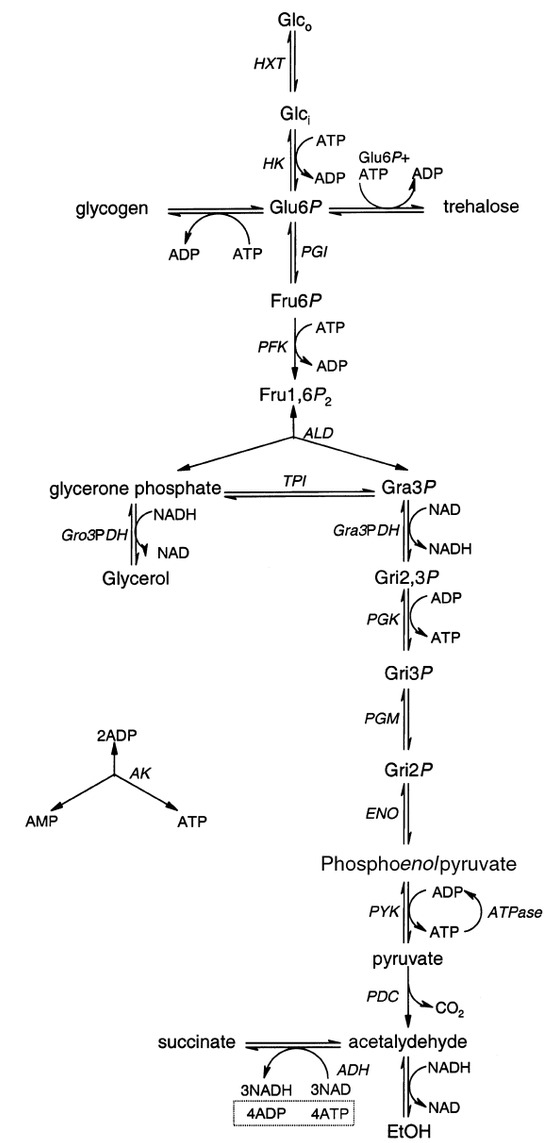


Fig. 1. Schematic of the model yeast glycolytic pathway. The boxed areas indicate the 'perturbation' of including ATP/ADP conversion in the succinate step which was present in [7], but not the provided SCAMP model. The ATP-ADP conversion is not included in the GEPASI model described herein. AMP, adenosine monophosphate; EtOH, ethanol; Fru1,6P_2, fructose 1,6-biphosphate; Fru6*P*, fructose 6-phosphate; GLCi, glucose (internal); GLCo, glucose (external); Gri3*P*, 3-phosphoglycerate; Gri3*P*, 2-phosphoglycerate; Gri1,3P_2, 1,3-bisphosphoglycerate; ADH, alcohol dehydrogenase; AK, adenylate kinase; ALD, aldolase; ENO, enolase; Gro3*P*DH, glycerol-3-phosphate dehydrogenase; Glc6*P*, glucose-6-phosphate; Gra3*P*, D-glyceraldehyde-3-phosphate; Gra3*P*DH, glyceraldehyde-3-phosphate dehydrogenase; HK, hexokinase; HXT, hexose transport; PDC, pyruvate decarboxylase; PFK; phosphofructokinase; PGI, phosphoglucoisomerase; PGK, phosphoglycerate kinase; PGM, phosphoglycerate mutase; PYK, pyruvate kinase; TPI, triosephosphate isomerase.

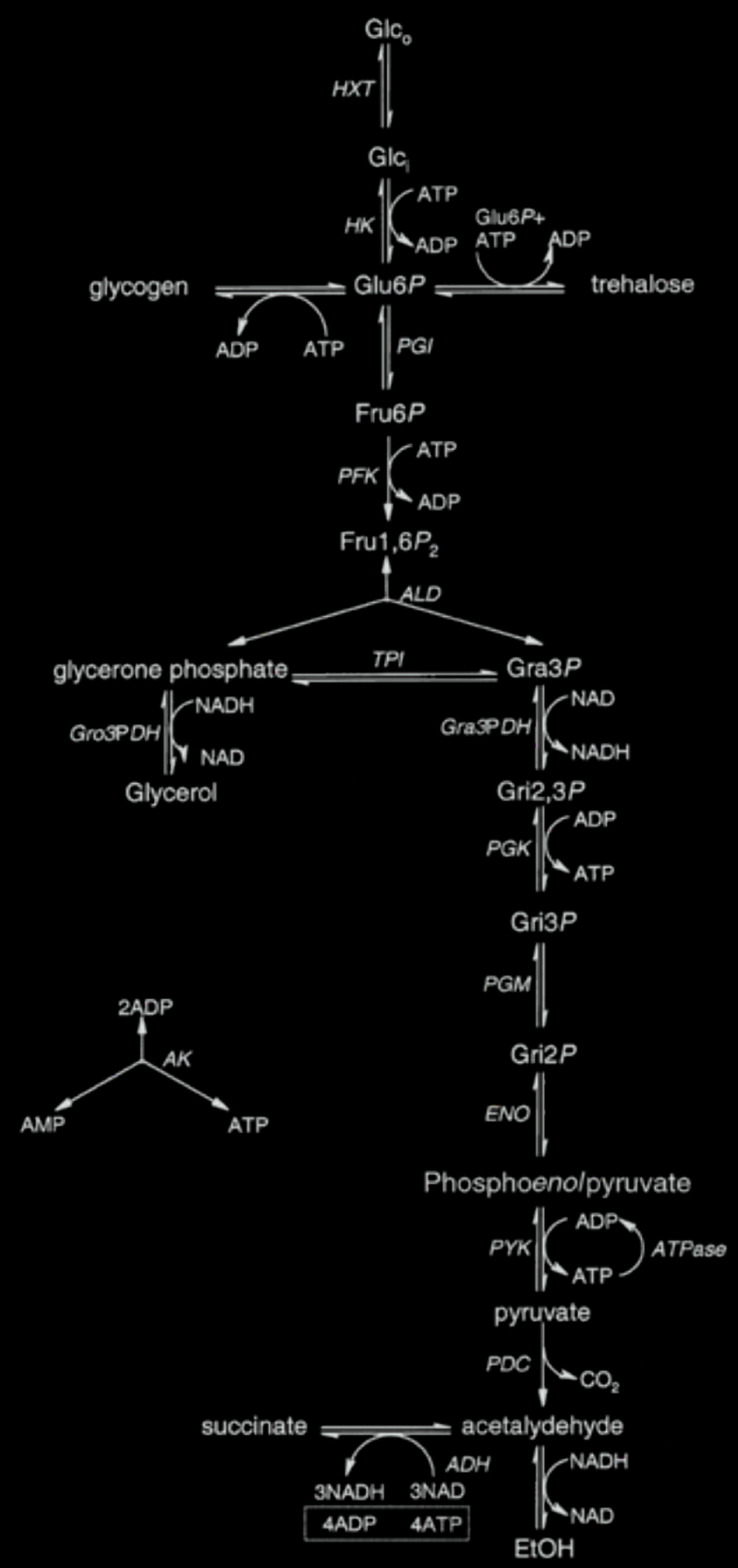
Correspondence to D. B. Kell, Cledwyn Building, Institute of Biological Sciences, University of Wales, Aberystwyth, Wales, UK, SY23 3DD.
Fax: + 44 1970622354, Tel.: + 44 1970622334,
E-mail: dbk@aber.ac.uk

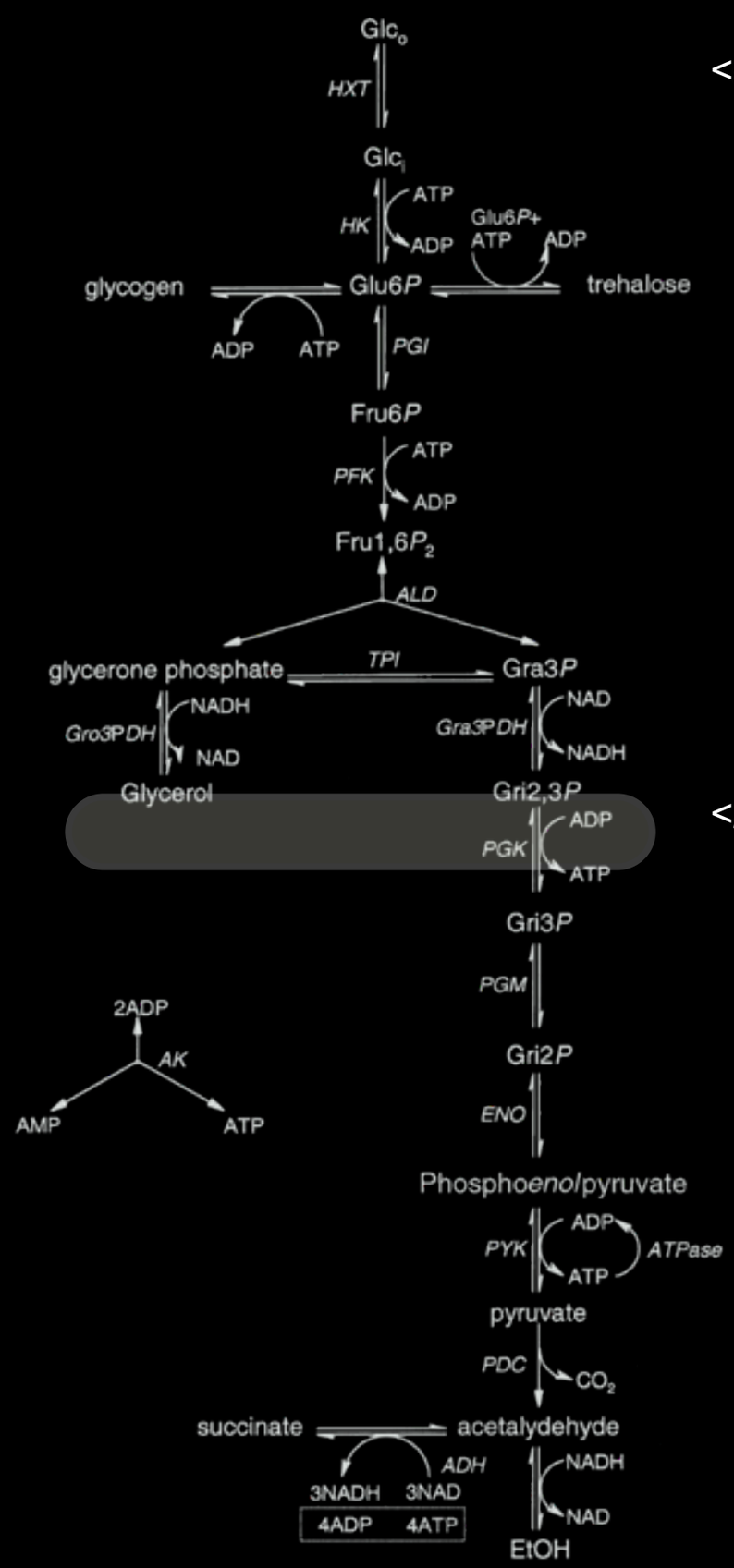
Abbreviations: PCA, principal components analysis; C^J_X, flux control coefficient for step X and flux J; FCC, flux control coefficient; Glyc, glycogen branch; Succ, succinate branch; Treh, trehalose branch.

Enzymes: alcohol dehydrogenase (EC 1.1.1.1); adenylate kinase (EC 2.7.4.2); aldolase (EC 4.1.2.13); enolase (EC 4.2.1.11); glycerol-3-phosphate dehydrogenase (EC 1.1.99.5); glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12); hexokinase (EC 2.7.1.1); pyruvate decarboxylase (EC 4.1.1.1); phosphofructokinase (EC 2.7.1.11); phosphoglucoisomerase (EC 5.3.1.9); phosphoglycerate kinase (EC 2.7.2.3); phosphoglycerate mutase (EC 5.4.2.1); pyruvate kinase (EC 2.7.1.40); triosephosphate isomerase (EC 5.3.1.1).

Note: a web site is available at http://qbab.dbs.aber.ac.uk

(Received 24 January 2002, revised 10 June 2002, accepted 18 June 2002)


Glc$_o$
HXT
Glc$_i$
ATP
HK
ADP
Glu6P+
ATP
ADP
glycogen
Glu6P
trehalose
ADP
ATP
PGI
Fru6P
ATP
PFK
ADP
Fru1,6P$_2$
ALD
glycerone phosphate
TPI
Gra3P
NADH
Gro3P DH
NAD
NAD
Gra3P DH
NADH
Glycerol
Gri2,3P
ADP
PGK
ATP
Gri3P
PGM
2ADP
AK
AMP
ATP
Gri2P
ENO
Phosphoenol/pyruvate
ADP
PYK
ATP
ATPase
pyruvate
PDC
CO$_2$
succinate
acetalydehyde
ADH
3NADH
3NAD
4ADP
4ATP
NADH
NAD
EtOH




```
<species metaid="metaid_0000012" id="PEP" name="Phosphoenolpyruvate"
        compartment="cytosol" initialConcentration="0.07">
      <annotation>
        <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
                 xmlns:bqbiol="http://biomodels.net/biology-qualifiers/"
                 xmlns:bqmodel="http://biomodels.net/model-qualifiers/" >
          <rdf:Description rdf:about="#metaid_0000012">
              <bqbiol:is>
                <rdf:Bag>
                  <rdf:li rdf:resource="http://www.ebi.ac.uk/chebi/#CHEBI:26055"/>
                  <rdf:li rdf:resource="http://www.genome.jp/kegg/compound/#C00074"/>
                </rdf:Bag>
              </bqbiol:is>
          </rdf:Description>
        </rdf:RDF>
      </annotation>
</species>
```

equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

[1] Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, **40**, 149 (1920).
[2] Longuet-Higgins, M. S., *Mon. Not. Roy. Astro. Soc., Geophys. Supp.*, **5**, 285 (1949).
[3] Von Arx, W. S., Woods Hole Papers in Phys. Oceanog. Meteor., **11** (3) (1950).
[4] Ekman, V. W., *Arkiv. Mat. Astron. Fysik. (Stockholm)*, **2** (11) (1905).

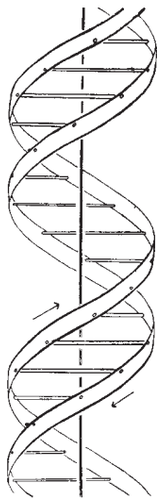
MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.



This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3′,5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the *z*-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical *z*-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at

King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J. D. WATSON
F. H. C. CRICK

Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems,
Cavendish Laboratory, Cambridge.
April 2.

[1] Pauling, L., and Corey, R. B., *Nature*, **171**, 346 (1953); *Proc. U.S. Nat. Acad. Sci.*, **39**, 84 (1953).
[2] Furberg, S., *Acta Chem. Scand.*, **6**, 634 (1952).
[3] Chargaff, E., for references see Zamenhof, S., Brawerman, G., and Chargaff, E., *Biochim. et Biophys. Acta*, **9**, 402 (1952).
[4] Wyatt, G. R., *J. Gen. Physiol.*, **36**, 201 (1952).
[5] Astbury, W. T., Symp. Soc. Exp. Biol. 1, Nucleic Acid, 66 (Camb. Univ. Press, 1947).
[6] Wilkins, M. H. F., and Randall, J. T., *Biochim. et Biophys. Acta*, **10**, 192 (1953).

Molecular Structure of Deoxypentose Nucleic Acids

WHILE the biological properties of deoxypentose nucleic acid suggest a molecular structure containing great complexity, X-ray diffraction studies described here (cf. Astbury[1]) show the basic molecular configuration has great simplicity. The purpose of this communication is to describe, in a preliminary way, some of the experimental evidence for the polynucleotide chain configuration being helical, and existing in this form when in the natural state. A fuller account of the work will be published shortly.

The structure of deoxypentose nucleic acid is the same in all species (although the nitrogen base ratios alter considerably) in nucleoprotein, extracted or in cells, and in purified nucleate. The same linear group of polynucleotide chains may pack together parallel in different ways to give crystalline[1–3], semi-crystalline or paracrystalline material. In all cases the X-ray diffraction photograph consists of two regions, one determined largely by the regular spacing of nucleotides along the chain, and the other by the longer spacings of the chain configuration. The sequence of different nitrogen bases along the chain is not made visible.

Oriented paracrystalline deoxypentose nucleic acid ('structure *B*' in the following communication by Franklin and Gosling) gives a fibre diagram as shown in Fig. 1 (cf. ref. 4). Astbury suggested that the strong 3·4-A. reflexion corresponded to the inter-nucleotide repeat along the fibre axis. The ~ 34 A. layer lines, however, are not due to a repeat of a polynucleotide composition, but to the chain configuration repeat, which causes strong diffraction as the nucleotide chains have higher density than the interstitial water. The absence of reflexions on or near the meridian immediately suggests a helical structure with axis parallel to fibre length.

Diffraction by Helices

It may be shown[5] (also Stokes, unpublished) that the intensity distribution in the diffraction pattern of a series of points equally spaced along a helix is given by the squares of Bessel functions. A uniform continuous helix gives a series of layer lines of spacing corresponding to the helix pitch, the intensity distribution along the *n*th layer line being proportional to the square of J_n, the *n*th order Bessel function. A straight line may be drawn approximately through

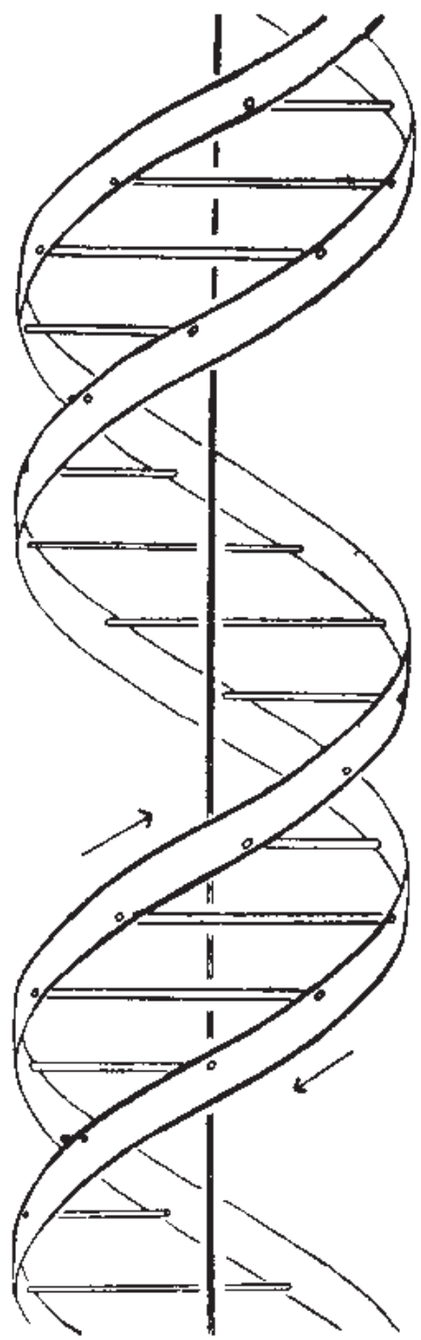
Fig. 1. Fibre diagram of deoxypentose nucleic acid from *B. coli*. Fibre axis vertical

the innermost maxima of each Bessel function and the origin. The angle this line makes with the equator is roughly equal to the angle between an element of the helix and the helix axis. If a unit repeats *n* times along the helix there will be a meridional reflexion (J_0^2) on the *n*th layer line. The helical configuration produces side-bands on this fundamental frequency, the effect[5] being to reproduce the intensity distribution about the origin around the new origin, on the *n*th layer line, corresponding to *C* in Fig. 2.

We will now briefly analyse in physical terms some of the effects of the shape and size of the repeat unit or nucleotide on the diffraction pattern. First, if the nucleotide consists of a unit having circular symmetry about an axis parallel to the helix axis, the whole diffraction pattern is modified by the form factor of the nucleotide. Second, if the nucleotide consists of a series of points on a radius at right-angles to the helix axis, the phases of radiation scattered by the helices of different diameter passing through each point are the same. Summation of the corresponding Bessel functions gives reinforcement for the inner-

Fig. 2. Diffraction pattern of system of helices corresponding to structure of deoxypentose nucleic acid. The squares of Bessel functions are plotted about O on the equator and on the first, second, third and fifth layer lines for half of the nucleotide mass at 20 A. diameter and remainder distributed along a radius, the mass at a given radius being proportional to the radius. About *C* on the tenth layer line similar functions are plotted for an outer diameter of 12 A.

structure as described is rather ill-defined, and for this reason we shall not comment on it.



This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3′,5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There

It has been found experimental
of the amounts of adenine to thym
of guanine to cytosine, are always
for deoxyribose nucleic acid.

It is probably impossible to bu
with a ribose sugar in place of th
the extra oxygen atom would mak
der Waals contact.

The previously published X-ray
ribose nucleic acid are insufficient
of our structure. So far as we can
compatible with the experimental
be regarded as unproved until it
against more exact results. Some
in the following communications.
of the details of the results presen
devised our structure, which rests
entirely on published experimenta
chemical arguments.

It has not escaped our notice
pairing we have postulated imme
possible copying mechanism for th

Full details of the structure, i
ditions assumed in building it, to
of co-ordinates for the atoms,
elsewhere.

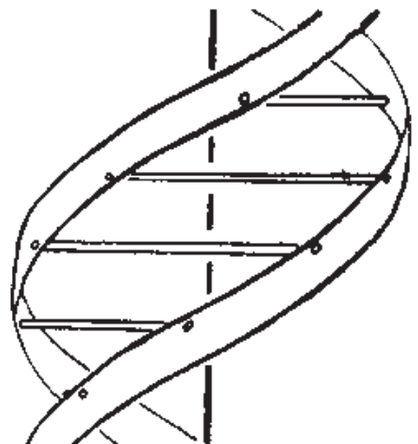
We are much indebted to Dr.
constant advice and criticism, es
atomic distances. We have also b
a knowledge of the general nature
experimental results and ideas
Wilkins, Dr. R. E. Franklin and t

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.



We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical

hydrogen-bonded to a single base
chain, so that the two lie side by si
z-co-ordinates. One of the pair mus
the other a pyrimidine for bondin
hydrogen bonds are made as follows
1 to pyrimidine position 1; puri
pyrimidine position 6.

If it is assumed that the bases
structure in the most plausible t
(that is, with the keto rather tha
figurations) it is found that only
bases can bond together. These pa
(purine) with thymine (pyrimidin
(purine) with cytosine (pyrimidine)

In other words, if an adenine forn
a pair, on either chain, then on t
the other member must be thymin
guanine and cytosine. The sequen
single chain does not appear to be
way. However, if only specific pair
formed, it follows that if the sequ
one chain is given, then the seque
chain is automatically determined.

It has been found experimentally
of the amounts of adenine to thymi
of guanine to cytosine, are always v
for deoxyribose nucleic acid.

It is probably impossible to bui
with a ribose sugar in place of th
the extra oxygen atom would mak
der Waals contact.

The previously published X-ray
ribose nucleic acid are insufficient f

m). We
chemical
hat each
hate di-
o-deoxy-
ith 3′,5′
ins (but
ted by a
the fibre
w right-
wing to
s of the
ins run
Each
es Fur-
that is,
nside of
hates on
guration
e atoms
urberg's
on', the
erpendi-

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F.

Ceci n'est pas un hamburger.

equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

[1] Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, **40**, 149 (1920).
[2] Longuet-Higgins, M. S., *Mon. Not. Roy. Astro. Soc., Geophys. Supp.*, **5**, 285 (1949).
[3] Von Arx, W. S., Woods Hole Papers in Phys. Oceanog. Meteor., **11** (3) (1950).
[4] Ekman, V. W., *Arkiv. Mat. Astron. Fysik. (Stockholm)*, **2** (11) (1905).

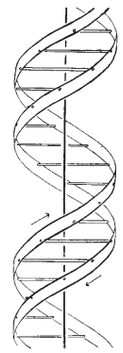
MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.



This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3′,5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the *z*-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical *z*-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at

Ceci n'est pas un hamburger.

Ceci n'est pas un hamburger.

Revolution or Evolution?

MANCHESTER
1824

The University
of Manchester

http://utopiadocs.com/

@utopiadocs @srp

Alice
Villeger

Dave
Thorne

James
Marsh

Steve
Pettifer

Philip
McDermott

Terri
Attwood

MANCHESTER 1824

The University of Manchester

Open PHACTS
Open Pharmacological Space

http://utopiadocs.com/

@utopiadocs

BBSRC bioscience for the future

EPSRC Pioneering research and skills

portlandpresslimited publishing innovation

MBRACE Grid Network of excellence

EasyChair conference system

Pfizer

ELSEVIER

AstraZeneca

Ceci n'est pas un hamburger.

doi:10.1087/20110309