by 31.9%. When the figures are tested statistically by the χ^2 method, it is found that this is a significant difference, the probability of its occurring as a result of chance being less than 0.01.

There was considerable variation in the size of the tumors. In some instances, the entire lobe of the lung was a mass of tumorous tissue (Fig. 2). In others, the

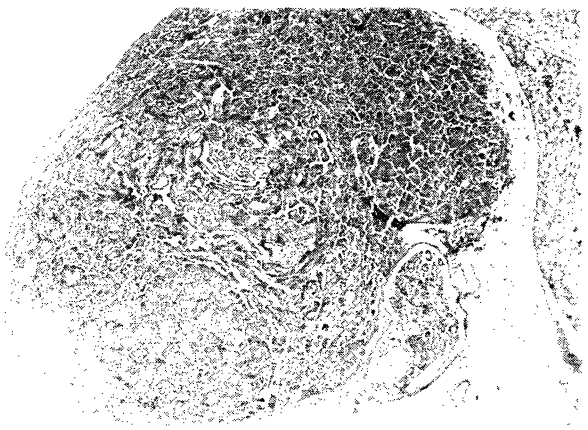

FIG. 2. Primary neoplasm of the lung of a mouse exposed to cigarette smoke for 1 year. It is of unifocal origin and involved practically the whole lobe of the lung. × 50.

tumor was of small size. In some of the lungs growth started from a single focus, part of a single bronchus. In others, it was multifocal (Fig. 3). Papillary adeno-

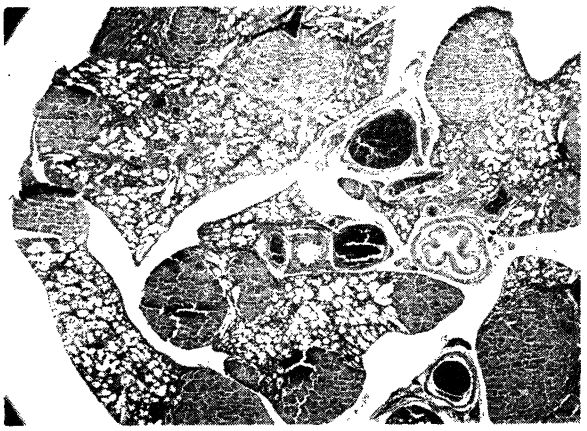

FIG. 3. Primary neoplasm of multifocal origin in the lung of a mouse exposed to cigarette smoke for 14 months. × 25.

carcinoma was the most common type of tumor, but adenomas and adenocarcinomas were also found. Lymphoid infiltration was a common occurrence; masses of lymph cells often surrounded branches of the bronchial tree and blood vessels. In some instances, lymphoid accumulation filled the greater part of a lobe of the lung. Pathology was noted in the organs of reproduction, endocrine glands, the kidney, and the liver.

The weight records showed that the smoked mice grew more slowly and failed by a large margin to attain the weight of the controls. Although sexes were equally divided in the second experiment, and there were some males in the first experiment, no young were obtained from the experimental mice. The controls, on the contrary, reproduced freely.

The smoked mice of the two experiments can be lumped together because they were exposed to cigarette smoke for at least one year and because the results of both experiments were similar, if not identical.

Since no other experimental differences existed between the smoked and the control mice except the smoke, it seems justifiable to conclude that the preponderance of tumors in the experimental mice was induced by the cigarette smoke. Cigarette smoke consists of many ingredients, some of which are considered carcinogenic. Of these, the tars and arsenic (*2*) are the foremost. There is some reason to believe that the alkaloids of nicotine have irritating properties that may act in the production of tumors. Which of these alone or in combination is responsible for the production of tumors in mice and possibly in other animals remains to be proved.

The weight and growth rates of the experimental animals were appreciably lower than those for the controls, but no explanation is available for this phenomenon at the present time. It could be due either to an inadequate intake of food or to the decreased utilization of food.

The lack of reproduction among the experimental animals is known to be caused by atrophic changes of the reproductive system (*3*). It is further known that the pathology caused by the injection of nicotine solution into mice or rats will parallel the pathology of the reproductive organs of mice exposed to cigarette smoke (*3*).

References

1. LORENZ, E., *et al. Cancer Research*, **3**, 123 (1943).
2. THOMAS, M. D., and COLLIER, T. R. *J. Ind. Hyg. Toxicol.*, **27**, 201 (1945).
3. ESSENBERG, J. M., FAGAN, L., and MALERSTEIN, A. J. *Western J. Surg. Obstet., Gynecol.*, **59**, 27 (1951).

Manuscript received May 14, 1952.

Molecules of the Insulin Structure[1]

Dorothy Wrinch

Department of Physics, Smith College, Northampton, Massachusetts

Structures suggested for skeletons of protein molecules in 1948 (*1–3*) have become of greater interest in the light of results regarding insulin recently obtained.

The original C_n models (the skeletons of the cage-like, space-enclosing polycondensations of α-amino acid molecules proposed [*4*] as models for protein molecules in 1936) comprise residues interlocked at both terminals, with every atom of the NC_aC backbones (belonging to the constituent amino acid molecules $NH_2—C_aHR—COOH$) on or near a point of

[1] This work is supported by the Office of Naval Research.

a diamond network. The presence of this underlying regular network has the advantage of explaining how a structure comprising many atoms can yet achieve coherence. Further, since the valency angles in the diamond network are the tetrahedral angle, this coherence can be achieved with tetrahedral valency angles for the carbons and the nitrogens as would seem to be required. Once these basic aspects of the "diamond" cage hypothesis are appreciated, it becomes of interest to consider whether any structures comprising fewer than the 72 residues of the original C_1 structure can build a coherent whole. It is at once evident (*1*) that as many as one third of all the residues can be deleted without impairing the coherence. Thus we may retain both tetrahedral sets of cyclic hexapeptides, discarding the 6 cyclic tetrapeptides, or we may discard one set and retain the 6 tetrapeptides, obtaining in each case a 48-residue skeleton. Coherent structures with numbers of skeletal residues between 48 and 72 can also be obtained, in considerable variety, with any of the point group symmetries T, D_2, C_3, C_2, or C_1. A model of the structure of (tetrahedral) symmetry T comprising only the 4 + 4 cyclic hexapeptides is shown in Fig. 1. As may be seen, the spatial patterns of the various CO's, NH's, C(OH)'s, and N's are propitious for the formation of hydrogen bridge systems over the surface. There is correspondingly a T structure comprising one set of 4 cyclic hexapeptides and the set of 6 tetrapeptides. Further, structures with symmetry D_2 or C_3 or C_2 or C_1 can be visualized in which not all the residues of the 6 tetrapeptides and not all the residues of a quartette of cyclic hexapeptides are deleted.

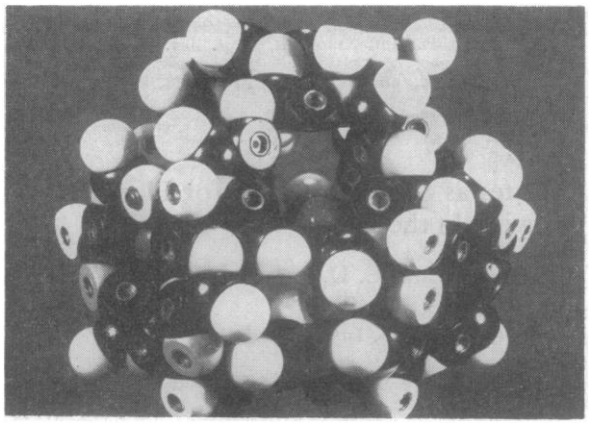

FIG. 1. A model of the 48-residue skeleton with symmetry T comprising two tetrahedral sets of cyclic hexapeptides, with half the residues in the classical form and half in the double peptide form.

Every deletion of a residue from the 72-residue structure leaves a "raw edge," with a residue in the classical form CO—CHR—NH, and in both the 48-residue structures described there are 24 in this form and 24 in the double peptide form C(OH)—CHR—N. At any CO or NH of a raw edge, a residue—or indeed a linear peptide—can be inserted by the appropriate single terminal, $—NH_2$ or —COOH, either by the change, $—NH_2$, —CO— to —(—NH)C(OH)—, or by the elimination of water, —COOH, —NH— to —(—CO)N—. In both cases the residues so inserted are functioning simply as substituents (*1, 3*).

We may therefore visualize the possibility of protein molecules with a number of skeletal residues as small as 48 to which there may, but need not, be added (1) more skeletal residues and (2) residues (or linear peptides) inserted in the skeleton at only one terminal, the number with free α-amino terminals not necessarily being equal to the number with free α-carboxyl terminals. The residue number for the structure is then the sum of the skeletal residues plus the substituent residues.

Information regarding the proportions of residues with free α-amino groups is now available for a number of proteins (*5–7*). It is never large, the proportion of 4/1200 molecular weight for insulin marking a relatively high content (*8*); in at least two cases, egg albumin (*9*) and tropomyosin (*10*), there are none. Corresponding information regarding residues with free α-carboxyl groups is beginning to appear (*11*).

For these new cage structures, the points of interest are, then, as follows: The models permit molecules with as few as 48 skeletal residues and provide sites for substituent residues; such structures can have the same or different numbers of residues with free α-amino groups and residues with free α-carboxyl groups, and either of these numbers can be 0; for such structures molecular weights of the order of, say, $48 \times 115 = 5520$ and up are to be expected.

Insulin. If, leaving aside certain difficulties in harmonizing various findings regarding the molecular weight of insulin, which have recently been discussed in some detail (*12–14*), we work simply from the crystallographic data, we begin (*15*) with an estimated molecular weight of *ca* 35,700 for the insulin particle P. Since the crystal is trigonal, it follows that $P = 3(Q_1 + Q_2 + \ldots + Q_m) + R_1 + R_2 \ldots R_n$, where the R's are trigonal and no symmetry is to be assumed for the Q's and m or n can be 0. In 1948 Oncley and Ellenbogen (*12, 16, 17*) observed the dissociation of this structure into "thirds" at sufficiently acid pH. They succeeded in crystallizing these subunits in the absence of zinc in the presence of sulfuric acid at pH 2–3.5. That these structures are, indeed, actual thirds follows from a preliminary x-ray study of this crystal by Low (*18*). From unit cell dimensions and a postulated density of 1.3, a molecular weight of 13,500 is deduced (*18*). Since this is to cover a 5% complement of water by weight and 12 sulfate groups per structure, the mol wt for the subunit comes out at *ca* 11,700, differing little from Chibnall's estimate (*19*) of 12,000 on chemical grounds to which he assigns a residue number of 106. Since the researches of both Brand (*20*) and Chibnall (*19*) yield an estimated average residue weight for insulin of about 114, a mol wt of 11,700 would correspond to about 102 residues. In the light

of these findings, there are no R's in the insulin particle and the specification of the subunit may be written to the form $\frac{1}{3}\ P = Q_1 + Q_2 \ldots Q_m$, where m need not be unity, with, say, about 102 residues to be distributed among all the Q's. To this complement correspond two of Sanger's A and B components (*5*). It will thus be necessary, ultimately, to prove, for any proposed structure of insulin, that these 4 peptide chains could be obtained from it by Sanger's procedures.

Meanwhile, in the light of early general studies of the language of the vector space (*21*), it was becoming clear that the vector map for the wet insulin crystal calculated from Crowfoot's intensities could be interpreted, in the neighborhood of the origin, in terms of cage structures of about the tetrahedral dimensions of the C_1 structures (*22*), which had already been shown to maintain their coherence with as few as 48 skeletal residues. On this basis, it was suggested (*22*) that each subunit of insulin may perhaps correspond to two sub-sub-units of this type, either as separate entities or (since at that time dissociation below thirds had not been observed) as an "intergrowth." Naturally nothing could be presumed as to whether the two cages are identical in the nature and placing of every residue or as to whether the actual residue numbers, or even the skeletal residue numbers, $48 + k$, $48 + k'$, are the same, or as to which of the residues of the original 72-residue skeleton are deleted.

The two latest developments are therefore of great interest. On the basis of a physicochemical investigation, Fredericq and Neurath (*23*) claim that the insulin particle is dissociable into structures of mol wt *ca* 6000, a result which was strongly suggested by the picture in terms of two cages for each subunit, though the possibility that these form an intergrowth was mentioned, in default at that time, of just this evidence. It seems questionable whether we may infer, even from the finding of a solubility curve indicative of a single component (*23*), that these two "6000" members corresponding to a subunit are identical in the nature and placing of every residue. That this may not be so is, in fact, suggested by the still more recent work of Harfenist and Craig (*24*).

It is not feasible at this time to draw final conclusions from the experimental data on insulin, which are of varied types, with margins of error not precisely specifiable and certain discrepancies still unresolved. In such a case, it is often of value to proceed on a different tack and study, as in this communication, how far the data can be fitted into a model suggested on structural grounds. Actually, the cumulative evidence seems to fit as well as it is reasonable to expect with the picture given above, which analyzes each of the 3 subunits of the insulin structure into two structures, a Q_1 and a Q_2, separate or as an intergrowth, and corresponding together to two of Sanger's A components and two of his B components. Taking the number of residues to be, say, 102, we have for each Q the necessary minimum of 48 skeletal residues.

References

1. Wrinch, D. *Science,* **107**, 445 (1948).
2. ———. *Biol. Bull.,* **95**, 247 (1948).
3. ———. *Wallerstein Lab. Communs.,* **11**, 175 (1948).
4. ———. *Nature,* **139**, 972 (1937).
5. Sanger, F. *Advances in Protein Chem.,* **7**, 2 (1952).
6. Porter, R. R., and Sanger, F. *Biochem. J.,* **42**, 287 (1948).
7. Porter, R. R. *Biochem. et Biophys. Acta,* **2**, 105 (1948).
8. Chibnall, A. C. *Proc. Roy. Soc.* (London), **B131**, 136 (1942).
9. Porter, R. R. Cited by F. Sanger in *Biochemical Symp. No. 3.* New York: Cambridge Univ. Press, 29 (1950).
10. Bailey, K. *Biochem. J.,* **49**, 23 (1951).
11. Fromageot, C. *Biochem. et Biophys. Acta,* **6**, 283 (1950).
12. Ellenbogen, E. Ph.D. thesis, Harvard Univ. (1949).
13. Pedersen, K. O. *Cold Spring Harbor Symposia Quant. Biol.,* **14**, 148 (1949).
14. Fromageot, C. *Ibid.,* 49.
15. Crowfoot, D. *Ibid.,* 72.
16. Oncley, J. L. Private communication (1948).
17. Oncley, J. L., and Ellenbogen, E. Paper presented at 116th mtg. Am. Chem. Soc. (1949).
18. Low, B. *Cold Spring Harbor Symposia Quant. Biol.,* **14**, 77 (1949).
19. Chibnall, A. C. *J. Intern. Soc. Leather Trades' Chemists,* **30**, 1 (1946).
20. Brand, E. *Ann. N. Y. Acad. Sci.,* **47**, 187 (1946).
21. Wrinch, D. *Phil. Mag.,* **30**, 98, 490 (1939).
22. ———. *J. Chem. Phys.,* **16**, 1007 (1948).
23. Fredericq, E., and Neurath, H. *J. Am. Chem. Soc.,* **72**, 2684 (1950).
24. Harfenist, E., and Craig, L. C. *Ibid.,* **73**, 877 (1951).

Manuscript received July 10, 1952.

Chelates as Sources of Iron for Plants Growing in the Field[1]

Ivan Stewart and C. D. Leonard

Citrus Experiment Station,
University of Florida, Lake Alfred

Iron deficiency is widely recognized as one of the most important problems in plant nutrition. It is known to occur in most of the major fruit-producing areas of the world. The symptom of iron deficiency is similar in most plants in that the interveinal areas of the leaves become chlorotic while the veins remain green (Fig. 1). In citrus the leaves often turn an ivory color and usually drop, many of the branches die, and production of the trees is greatly reduced.

Iron deficiency is found under two widely different sets of conditions. In the first place, it occurs extensively in crops growing on calcareous soils, where it is referred to as lime-induced chlorosis. Second, in certain regions, as in Florida, iron deficiency occurs most commonly in crops growing on acid soils. The presence of iron chlorosis does not always mean, however, that there is a shortage of iron in the soil. Even in the very light sandy soils of Florida there is usually sufficient iron for plant growth, provided this element can be utilized effectively.

It has been reported by various workers (*1–3*) that iron chlorosis can be induced by an excessive accumulation in the soil of heavy metals such as copper, manganese, zinc, nickel, or cobalt. Hewitt (*1*) has listed these elements according to their increasing ability to induce iron chlorosis in plants. The same worker makes

[1] Florida Agricultural Experiment Station Journal Series, No. 74.