

= Cyclol =

The cyclol hypothesis is the first structural model of a folded, globular protein. It was developed by Dorothy Wrinch in the late 1930s, and was based on three assumptions. Firstly, the hypothesis assumes that two peptide groups can be crosslinked by a cyclol reaction (Figure 1); these crosslinks are covalent analogs of non-covalent hydrogen bonds between peptide groups. These reactions have been observed in the ergopeptides and other compounds. Secondly, it assumes that, under some conditions, amino acids will naturally make the maximum possible number of cyclol crosslinks, resulting in cyclol molecules (Figure 2) and cyclol fabrics (Figure 3). These cyclol molecules and fabrics have never been observed. Finally, the hypothesis assumes that globular proteins have a tertiary structure corresponding to Platonic solids and semiregular polyhedra formed of cyclol fabrics with no free edges. Such "closed cyclol" molecules have not been observed either.

Although later data demonstrated that this original model for the structure of globular proteins needed to be amended, several elements of the cyclol model were verified, such as the cyclol reaction itself and the hypothesis that hydrophobic interactions are chiefly responsible for protein folding. The cyclol hypothesis stimulated many scientists to research questions in protein structure and chemistry, and was a precursor of the more accurate models hypothesized for the DNA double helix and protein secondary structure. The proposal and testing of the cyclol model also provides an excellent illustration of empirical falsifiability acting as part of the scientific method.

= = Historical context = =

By the mid-1930s, analytical ultracentrifugation studies by Theodor Svedberg had shown that proteins had a well-defined chemical structure, and were not aggregations of small molecules. The same studies appeared to show that the molecular weight of proteins fell into a few well-defined classes related by integers, such as  $M_w = 2p3q$  Da, where  $p$  and  $q$  are nonnegative integers. However, it was difficult to determine the exact molecular weight and number of amino acids in a protein. Svedberg had also shown that a change in solution conditions could cause a protein to disassemble into small subunits, now known as a change in quaternary structure.

The chemical structure of proteins was still under debate at that time. The most accepted (and ultimately correct) hypothesis was that proteins are linear polypeptides, i.e., unbranched polymers of amino acid residues, and several distinguished scientists were unsure whether such long, linear macromolecules could be stable in solution. Further doubts about the polypeptide nature of proteins arose because some enzymes were observed to cleave proteins but not peptides, whereas other enzymes cleave peptides but not folded proteins. Attempts to synthesize proteins in the test tube were unsuccessful, mainly due to the chirality of amino acids; naturally occurring proteins are composed of only left-handed amino acids. Hence, alternative chemical models of proteins were considered, such as the diketopiperazine hypothesis of Emil Abderhalden. However, no alternative model had yet explained why proteins yield only amino acids and peptides upon hydrolysis and proteolysis. As clarified by Linderstrøm-Lang, these proteolysis data showed that denatured proteins were polypeptides, but no data had yet been obtained about the structure of folded proteins; thus, denaturation could involve a chemical change that converted folded proteins into polypeptides.

The process of protein denaturation (as distinguished from coagulation) had been discovered in 1910 by Harriette Chick and Charles Martin, but its nature was still mysterious. Tim Anson and Alfred Mirsky had shown that denaturation was a reversible, two-state process that results in many chemical groups becoming available for chemical reactions, including cleavage by enzymes. In 1929, Hsien Wu hypothesized correctly that denaturation corresponded to protein unfolding, a purely conformational change that resulted in the exposure of amino acid side chains to the solvent. Wu's hypothesis was also advanced independently in 1936 by Mirsky and Linus Pauling. Nevertheless, protein scientists could not exclude the possibility that denaturation corresponded to

a chemical change in the protein structure , a hypothesis that was considered a ( distant ) possibility until the 1950s .

X-ray crystallography had just begun as a discipline in 1911 , and had advanced relatively rapidly from simple salt crystals to crystals of complex molecules such as cholesterol . However , even the smallest proteins have over 1000 atoms , which makes determining their structure far more complex . In 1934 , Dorothy Crowfoot Hodgkin had taken crystallographic data on the structure of the small protein , insulin , although the structure of that and other proteins were not solved until the late 1960s . However , pioneering X-ray fiber diffraction data had been collected in the early 1930s for many natural fibrous proteins such as wool and hair by William Astbury , who proposed rudimentary models of secondary structure elements such as the alpha helix and the beta sheet .

Since protein structure was so poorly understood in the 1930s , the physical interactions responsible for stabilizing that structure were likewise unknown . Astbury hypothesized that the structure of fibrous proteins was stabilized by hydrogen bonds in  $\beta$ -sheets . The idea that globular proteins are also stabilized by hydrogen bonds was proposed by Dorothy Jordan Lloyd in 1932 , and championed later by Alfred Mirsky and Linus Pauling . At a 1933 lecture by Astbury to the Oxford Junior Scientific Society , physicist Frederick Frank suggested that the fibrous protein  $\alpha$ -keratin might be stabilized by an alternative mechanism , namely , covalent crosslinking of the peptide bonds by the cyclol reaction above . The cyclol crosslink draws the two peptide groups close together ; the N and C atoms are separated by  $\sim 1.5 \text{ \AA}$  , whereas they are separated by  $\sim 3 \text{ \AA}$  in a typical hydrogen bond . The idea intrigued J. D. Bernal , who suggested it to the mathematician Dorothy Wrinch as possibly useful in understanding protein structure .

= = Basic theory = =

Wrinch developed this suggestion into a full fledged model of protein structure . The basic cyclol model was laid out in her first paper ( 1936 ) . She noted the possibility that polypeptides might cyclize to form closed rings ( true ) and that these rings might form internal crosslinks through the cyclol reaction ( also true , although rare ) . Assuming that the cyclol form of the peptide bond could be more stable than the amide form , Wrinch concluded that certain cyclic peptides would naturally make the maximal number of cyclol bonds ( such as cyclol 6 , Figure 2 ) . Such cyclol molecules would have hexagonal symmetry , if the chemical bonds were taken as having the same length , roughly  $1.5 \text{ \AA}$  ; for comparison , the N-C and C-C bonds have the lengths  $1.42 \text{ \AA}$  and  $1.54 \text{ \AA}$  , respectively .

These rings can be extended indefinitely to form a cyclol fabric ( Figure 3 ) . Such fabrics exhibit a long range , quasi-crystalline order that Wrinch felt was likely in proteins , since they must pack hundreds of residues densely . Another interesting feature of such molecules and fabrics is that their amino acid side chains point axially upwards from only one face ; the opposite face has no side chains . Thus , one face is completely independent of the primary sequence of the peptide , which Wrinch conjectured might account for sequence independent properties of proteins .

In her initial article , Wrinch stated clearly that the cyclol model was merely a working hypothesis , a potentially valid model of proteins that would have to be checked . Her goals in this article and its successors were to propose a well defined testable model , to work out the consequences of its assumptions and to make predictions that could be tested experimentally . In these goals , she succeeded ; however , within a few years , experiments and further modeling showed that the cyclol hypothesis was untenable as a model for globular proteins .

= = Stabilizing energies = =

In two tandem Letters to the Editor ( 1936 ) , Wrinch and Frank addressed the question of whether the cyclol form of the peptide group was indeed more stable than the amide form . A relatively simple calculation showed that the cyclol form is significantly less stable than the amide form . Therefore , the cyclol model would have to be abandoned unless a compensating source of energy

could be identified . Initially , Frank proposed that the cyclol form might be stabilized by better interactions with the surrounding solvent ; later , Wrinch and Irving Langmuir hypothesized that hydrophobic association of nonpolar sidechains provides stabilizing energy to overcome the energetic cost of the cyclol reactions .

The lability of the cyclol bond was seen as an advantage of the model , since it provided a natural explanation for the properties of denaturation ; reversion of cyclol bonds to their more stable amide form would open up the structure and allows those bonds to be attacked by proteases , consistent with experiment . Early studies showed that proteins denatured by pressure are often in a different state than the same proteins denatured by high temperature , which was interpreted as possibly supporting the cyclol model of denaturation .

The Langmuir @-@ Wrinch hypothesis of hydrophobic stabilization shared in the downfall of the cyclol model , owing mainly to the influence of Linus Pauling , who favored the hypothesis that protein structure was stabilized by hydrogen bonds . Another twenty years had to pass before hydrophobic interactions were recognized as the chief driving force in protein folding .

= = Steric complementarity = =

In her third paper on cyclols ( 1936 ) , Wrinch noted that many " physiologically active " substances such as steroids are composed of fused hexagonal rings of carbon atoms and , thus , might be sterically complementary to the face of cyclol molecules without the amino @-@ acid side chains . Wrinch proposed that steric complementarity was one of chief factors in determining whether a small molecule would bind to a protein .

Wrinch speculated that proteins are responsible for the synthesis of all biological molecules . Noting that cells digest their proteins only under extreme starvation conditions , Wrinch further speculated that life could not exist without proteins .

= = Hybrid models = =

From the beginning , the cyclol reaction was considered as a covalent analog of the hydrogen bond . Therefore , it was natural to consider hybrid models with both types of bonds . This was the subject of Wrinch 's fourth paper on the cyclol model ( 1936 ) , written together with Dorothy Jordan Lloyd , who first proposed that globular proteins are stabilized by hydrogen bonds . A follow @-@ up paper was written in 1937 that referenced other researchers on hydrogen bonding in proteins , such as Maurice Loyal Huggins and Linus Pauling .

Wrinch also wrote a paper with William Astbury , noting the possibility of a keto @-@ enol isomerization of the  $>C=O$  and an amide carbonyl group  $>C=O$  , producing a crosslink  $>C@-@C(OH?)$  and again converting the oxygen to a hydroxyl group . Such reactions could yield five @-@ membered rings , whereas the classic cyclol hypothesis produces six @-@ membered rings . This keto @-@ enol crosslink hypothesis was not developed much further .

= = Space @-@ enclosing fabrics = =

In her fifth paper on cyclols ( 1937 ) , Wrinch identified the conditions under which two planar cyclol fabrics could be joined to make an angle between their planes while respecting the chemical bond angles . She identified a mathematical simplification , in which the non @-@ planar six @-@ membered rings of atoms can be represented by planar " median hexagon " s made from the midpoints of the chemical bonds . This " median hexagon " representation made it easy to see that the cyclol fabric planes can be joined correctly if the dihedral angle between the planes equals the tetrahedral bond angle  $\theta = \arccos( -1/3 ) \approx 109.47^\circ$  .

A large variety of closed polyhedra meeting this criterion can be constructed , of which the simplest are the truncated tetrahedron , the truncated octahedron , and the octahedron , which are Platonic solids or semiregular polyhedra . Considering the first series of " closed cyclols " ( those modeled on the truncated tetrahedron ) , Wrinch showed that their number of amino acids increased

quadratically as  $72n^2$ , where  $n$  is the index of the closed cyclol  $C_n$ . Thus, the  $C_1$  cyclol has 72 residues, the  $C_2$  cyclol has 288 residues, etc. Preliminary experimental support for this prediction came from Max Bergmann and Carl Niemann, whose amino acid analyses suggested that proteins were composed of integer multiples of 288 amino acid residues ( $n = 2$ ). More generally, the cyclol model of globular proteins accounted for the early analytical ultracentrifugation results of Theodor Svedberg, which suggested that the molecular weights of proteins fell into a few classes related by integers.

The cyclol model was consistent with the general properties then attributed to folded proteins. (1) Centrifugation studies had shown that folded proteins were significantly denser than water ( $\sim 1.4$  g/mL) and, thus, tightly packed; Wrinch assumed that dense packing should imply regular packing. (2) Despite their large size, some proteins crystallize readily into symmetric crystals, consistent with the idea of symmetric faces that match up upon association. (3) Proteins bind metal ions; since metal binding sites must have specific bond geometries (e.g., octahedral), it was plausible to assume that the entire protein also had similarly crystalline geometry. (4) As described above, the cyclol model provided a simple chemical explanation of denaturation and the difficulty of cleaving folded proteins with proteases. (5) Proteins were assumed to be responsible for the synthesis of all biological molecules, including other proteins. Wrinch noted that a fixed, uniform structure would be useful for proteins in templating their own synthesis, analogous to the Watson and Francis Crick concept of DNA templating its own replication. Given that many biological molecules such as sugars and sterols have a hexagonal structure, it was plausible to assume that their synthesizing proteins likewise had a hexagonal structure. Wrinch summarized her model and the supporting molecular weight experimental data in three review articles.

== Predicted protein structures ==

Having proposed a model of globular proteins, Wrinch investigated whether it was consistent with the available structural data. She hypothesized that bovine tuberculin protein (523) was a  $C_1$  closed cyclol consisting of 72 residues and that the digestive enzyme pepsin was a  $C_2$  closed cyclol of 288 residues. These residue number predictions were difficult to verify, since the methods then available to measure the mass of proteins were inaccurate, such as analytical ultracentrifugation and chemical methods.

Wrinch also predicted that insulin was a  $C_2$  closed cyclol consisting of 288 residues. Limited X-ray crystallographic data were available for insulin which Wrinch interpreted as "confirming" her model. However, this interpretation drew rather severe criticism for being premature. Careful studies of the Patterson diagrams of insulin taken by Dorothy Crowfoot Hodgkin showed that they were roughly consistent with the cyclol model; however, the agreement was not good enough to claim that the cyclol model was confirmed.

== Downfall ==

The cyclol fabric was shown to be implausible for several reasons. Hans Neurath and Henry Bull showed that the dense packing of side chains in the cyclol fabric was inconsistent with the experimental density observed in protein films. Maurice Huggins calculated that several non-bonded atoms of the cyclol fabric would approach more closely than allowed by their van der Waals radii; for example, the inner H and C atoms of the lacunae would be separated by only 1.68 Å (Figure 5). Haurowitz showed chemically that the outside of proteins could not have a large number of hydroxyl groups, a key prediction of the cyclol model, whereas Meyer and Hohenemser showed that cyclol condensations of amino acids did not exist even in minute quantities as a transition state. More general chemical arguments against the cyclol model were given by Bergmann and Niemann and by Neuberger. Infrared spectroscopic data showed that the number of carbonyl groups in a protein did not change upon hydrolysis, and that intact, folded proteins have a full complement of amide carbonyl groups; both observations contradict the cyclol hypothesis that

such carbonyls are converted to hydroxyl groups in folded proteins . Finally , proteins were known to contain proline in significant quantities ( typically 5 % ) ; since proline lacks the amide hydrogen and its nitrogen already forms three covalent bonds , proline seems incapable of the cyclol reaction and of being incorporated into a cyclol fabric . An encyclopedic summary of the chemical and structural evidence against the cyclol model was given by Pauling and Niemann . Moreover , a supporting piece of evidence ? the result that all proteins contain an integer multiple of 288 amino acid residues ? was likewise shown to be incorrect in 1939 .

Wrinch replied to the steric clash , free energy , chemical and residue number criticisms of the cyclol model . On steric clashes , she noted that small deformations of the bond angles and bond lengths would allow these steric clashes to be relieved , or at least reduced to a reasonable level . She noted that distances between non bonded groups within a single molecule can be shorter than expected from their van der Waals radii , e.g. , the 2.93 Å distance between methyl groups in hexamethylbenzene . Regarding the free energy penalty for the cyclol reaction , Wrinch disagreed with Pauling 's calculations and stated that too little was known of intramolecular energies to rule out the cyclol model on that basis alone . In reply to the chemical criticisms , Wrinch suggested that the model compounds and simple bimolecular reactions studied need not pertain to the cyclol model , and that steric hindrance may have prevented the surface hydroxyl groups from reacting . On the residue number criticism , Wrinch extended her model to allow for other numbers of residues . In particular , she produced a " minimal " closed cyclol of only 48 residues , and , on that ( incorrect ) basis , may have been the first to suggest that the insulin monomer had a molecular weight of roughly 6000 Da .

Therefore , she maintained that the cyclol model of globular proteins was still potentially viable and even proposed the cyclol fabric as a component of the cytoskeleton . However , most protein scientists ceased to believe in it and Wrinch turned her scientific attention to mathematical problems in X ray crystallography , to which she contributed significantly . One exception was physicist Gladys Anslow , Wrinch 's colleague at Smith College , who studied the ultraviolet absorption spectra of proteins and peptides in the 1940s and allowed for the possibility of cyclols in interpreting her results . As the sequence of insulin began to be determined by Frederick Sanger , Anslow published a three dimensional cyclol model with sidechains , based on the backbone of Wrinch 's 1948 " minimal cyclol " model .

= = Partial redemption = =

The downfall of the overall cyclol model generally led to a rejection of its elements ; one notable exception was J. D. Bernal 's short lived acceptance of the Langmuir Wrinch hypothesis that protein folding is driven by hydrophobic association . Nevertheless , cyclol bonds were identified in small , naturally occurring cyclic peptides in the 1950s .

Clarification of the modern terminology is appropriate . The classic cyclol reaction is the addition of the NH amine of a peptide group to the C = O carbonyl group of another ; the resulting compound is now called an azacyclol . By analogy , an oxacyclol is formed when an OH hydroxyl group is added to a peptidyl carbonyl group . Likewise , a thiacyclol is formed by adding an SH thiol moiety to a peptidyl carbonyl group .

The oxacyclol alkaloid ergotamine from the fungus *Claviceps purpurea* was the first identified cyclol . The cyclic depsipeptide serratomolide is also formed by an oxacyclol reaction . Chemically analogous cyclic thiacyclols have also been obtained . Classic azacyclols have been observed in small molecules and tripeptides . Peptides are naturally produced from the reversion of azacyclols , a key prediction of the cyclol model . Hundreds of cyclol molecules have now been identified , despite Linus Pauling 's calculation that such molecules should not exist because of their unfavorably high energy .

After a long hiatus during which she worked mainly on the mathematics of X ray crystallography , Wrinch responded to these discoveries with renewed enthusiasm for the cyclol model and its relevance in biochemistry . She also published two books describing the cyclol theory and small peptides in general .

= = Illustration of the scientific method = =

The cyclol model of protein structure is an example of empirical falsifiability acting as part of the scientific method . An original hypothesis is made that accounts for unexplained experimental observations ; the consequences of this hypothesis are worked out , leading to predictions that are tested by experiment . In this case , the key hypothesis was that the cyclol form of the peptide group could be favored over the amide form . This hypothesis led to the predictions of the cyclol @-@ 6 molecule and the cyclol fabric , which in turn suggested the model of semi @-@ regular polyhedra for globular proteins . A key testable prediction was that a folded protein 's carbonyl groups should be largely converted to hydroxyl groups ; however , spectroscopic and chemical experiments showed that this prediction was incorrect . The cyclol model also predicts a high lateral density of amino acids in folded proteins and in films that does not agree with experiment . Hence , the cyclol model could be rejected and the search begun for new hypotheses of protein structure , such as the models of the alpha helix proposed in the 1940s and 1950s .

It is sometimes argued that the cyclol hypothesis should never have been advanced , because of its a priori flaws , e.g. , its steric clashes , its inability to accommodate proline , and the high free energy disfavoring the cyclol reaction itself . Although such flaws rendered the cyclol hypothesis implausible , they did not make it impossible . The cyclol model was the first well @-@ defined structure proposed for globular proteins , and too little was then known of intramolecular forces and protein structure to reject it immediately . It neatly explained several general properties of proteins and accounted for then @-@ anomalous experimental observations . Although generally incorrect , some elements of the cyclol theory were eventually verified , such as the cyclol reactions and the role of hydrophobic interactions in protein folding . A useful comparison is the Bohr model of the hydrogen atom , which was considered implausible from its inception , even by its creator , yet led the way to the ultimately correct theory of quantum mechanics . Similarly , Linus Pauling proposed a well @-@ defined model of DNA that was likewise implausible yet thought @-@ provoking to other investigators .

Conversely , the cyclol model is an example of an incorrect scientific theory of great symmetry and beauty , two qualities that can be regarded as signs of " obviously true " scientific theories . For example , the Watson @-@ Crick double helix model of DNA is sometimes said to be " obvious " because of its plausible hydrogen bonding and symmetry ; nevertheless , other , less symmetrical structures of DNA are favored under different conditions . Similarly , the beautiful theory of general relativity was considered by Albert Einstein as not needing experimental verification ; yet even this theory will require revision for consistency with quantum field theory .