Liversidge Research Lecture

No. 21
1978

ELEGANCE IN MOLECULAR DESIGN: THE COPPER SITE OF PHOTOSYNTHETIC ELECTRON-TRANSFER PROTEIN

HANS C. FREEMAN

The Royal Society of New South Wales

Hans Charles Freeman

This photograph is reproduced by permission of the Australian Journal of Chemistry and the RACI

Liversidge Research Lecturer No. 21, 1978.

HANS CHARLES FREEMAN 1929-2008

Hans Charles Freeman was born on 26 May 1929 in Breslau, Germany. His family arrived in Sydney as refugees in 1938. After attending Double Bay Public School, he received his secondary education at Sydney Boys' High School where he was taught science by the 'now-legendary' teacher Leonard A. Basser. He proceeded to the University of Sydney, graduating B.Sc.(1st. Class Hons.), with the University Medal in 1950. While holding appointments as Teaching Fellow and Temporary Lecturer, he carried out research on the dipole moments of organic molecules in the gas state under the supervision of Professor R.J.W. Le Fèvre, and graduated M.Sc. in 1952. In 1952-3 he was a Rotary Foundation Fellow at the California Institute of Technology. There, at the suggestion of Linus Pauling, he began research in X-ray crystal structure analysis with one of the leading thinkers and strategists of the subject, Edward W. Hughes. In 1954 he returned to the University of Sydney as a Lecturer in Chemistry, and completion of the research begun in California led to the award of a Sydney Ph.D. in 1957. He was promoted to Senior Lecturer in 1959, and to Reader in 1964. In 1971 he was appointed to the Foundation Chair of Inorganic Chemistry at the University of Sydney, and served as Head of the School of Chemistry in 1975-6. From 1995 to 1997 he held an appointment as Professor of Chemistry, before retiring formally as Professor Emeritus in 1998. During periods of study leave he was a George Ellery Hale Foundation Fellow at the California Institute of Technology (1958-1959); Visiting Lecturer in Inorganic Chemistry at the University of Basel (1966); Guest Professor in Inorganic Chemistry and Biochemistry at University of Göteborg (1966); Visiting Professor in Molecular Biophysics at Yale University (1968); Guest Professor in Inorganic Chemistry at the University of Copenhagen (1977); and Visiting Scholar in Chemistry at Stanford University and the Stanford Synchrotron Radiation Laboratory (1983).

Much of his research has been concerned with X-ray crystal structure analysis of biologically important coordination compounds; early studies on model compounds for metal protein interactions led to work on metalloproteins. Important side-trips have dealt with metal complexes of chemotherapeutic interest, and with the use of neutron diffraction data to determine the structures of metal complexes and organic compounds. His crystallographic work, especially in the area of metalloproteins, has led to numerous collaborative studies using techniques other than X-ray crystallography - particularly electronic and vibrational spectroscopies, and electrochemistry. Since the early 1980s he has made a number of contributions to the development of methods using synchrotron radiation for the X-ray crystallography and X-ray absorption spectroscopy of metalloproteins.

Hans Freeman was Chairman of the Division of Metal-Organic Chemistry of the Royal Australian Chemical Institute in 1971-3, Foundation President of the Society for Crystallographers in Australia in 1976-7, and Chairman of the Australian National Committee for Crystallography in 1984-1992. He was the Program Chairman of the XIVth International Congress of Crystallography in Perth in 1987, and has served on numerous other national and international conference committees. He has been a

member of the editorial boards of the *Journal of Coordination Chemistry* (1971-1980), *Journal of Inorganic Biochemistry* (1979--, and Associate Editor 1997--), and the *Journal of Biological Inorganic Chemistry* (1996-7). Major commitments to scientific policy committees include the Australian Department of Industry, Technology and Commerce (International Science and Technology Advisory Committee, 1989-1992), the Australian Science and Technology Council (committees of inquiry into major research facilities, 1989-1990 and 1991-2), Australian National Beamline Facility (Board of Management, 1991-5), and the Australian Synchrotron Research Program (Policy and Review Board, 1996--).

Honours and Awards (Pre-2001)

1969 FRACI (Fellow of the Royal Australian Chemical Institute)

1978 Liversidge Research Lecture, Royal Society of New South Wales

1980 Burrows Medal and Award for Inorganic Chemistry, RACI

1984 Fellow Royal Society of Chemistry

1984 FAA

1999 Leighton Medal, RACI

Biographical Source

Personal communication.

Obituary: G. Jackson, J. Harrowfield, "Vale Alan Sargeson: a Great Scientist", *Chemistry in Australia*, 2009, **76** (7), 15-16.

G. Mitchell, "Emeritus Professor Hans Freeman (1929-2008)," *Chem News: Newsletter of the University of Sydney School of Chemistry*, 2009 (issue 14), 5.

Biographical Memoir: B. Bosnich "Alan McLeod Sargeson 1930-2008*", Historical Records of Australian Science*, 2011, **22**, 246-276.

Scientific Publications by H.C. Freeman

Between 1950 and 1999 H.C. Freeman published some 137 research papers and 9 reviews/book chapters, mainly on crystal structures of metal complexes.

ELEGANCE IN MOLECULAR DESIGN: THE COPPER SITE OF PHOTOSYNTHETIC ELECTRON-TRANSFER PROTEIN*

HANS C. FREEMAN

School of Chemistry, University of Sydney, N.S.W.

ABSTRACT. Plastocyanin is an intensely blue protein which is essential for photosynthesis in green leaves and in some algae. The blue colour is associated with the presence of a single copper atom in each molecule of the protein. In terms of the absorbance per copper atom, plastocyanin is about a hundred times as blue as 'normal' cupric compounds. In addition, the protein has an unusual electron spin resonance spectrum and an anomalously high redox potential. The combination of these properties occurs in some other copper-proteins but has not yet been mimicked in any model compound of low molecular weight.

The recent X-ray crystal structure analysis of plastocyanin has revealed a molecule ideally suited to the biological function which it performs. The nature of the copper site is such as to produce the high redox potential which is required for electron-transfer between plastocyanin and its neighbours in the photosynthetic chain. The location of the copper site in the protein molecule provides at least two reasonable electron-transfer pathways. The exterior of the molecule has distinctive features which suggest that the protein interacts in specific ways with its redox partners and/or its environment.

Introduction

I have the honour to be the present custodian of a copy of the Sydney University Calendar issued 100 years ago. This copy is particularly important. The top right-hand corner of the cover bears the signature of *Archibald Liversidge*.

Like all University Calendars, the version of 1878-9 includes a list of the Professors. Liversidge was Professor of Mineralogy, Lecturer in Geology, and Demonstrator in Chemistry. The Hon. John Smith, M.D., was Professor of Chemistry and Experimental Physics. A little while later the Hon. John Smith dropped one of his resonance forms and became simply Professor of Experimental Physics. Liversidge was transformed into the Professor of Chemistry. It is a pity that the increasing complexity of Science and the need for ever greater specialisation have reduced the opportunities for such moves. In these days, when even the adjectives which precede some Chairs of Chemistry are jealously guarded, there is much to admire in the multidisciplinary agility of Sydney's Professors of 100 years ago.

Two pages of the 1878-9 University Calendar are annotated in Liversidge's own hand. On a scale of course fees, the sum of "1 guinea per term" for Mineralogy has been crossed out, and "2 guineas" has been substituted. (Chemistry with Practical Chemistry was rated at 6 guineas.) And on the page of regulations for the conduct of examinations, Liversidge has underlined: "*and the Professor or Lecturer in the School must be satisfied with (the*

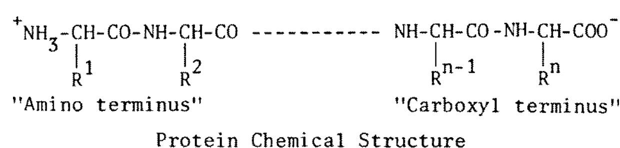
• The Liversidge Research Lecture, delivered before the Royal Society of New South Wales, 19th July, 1978. Reproduced by permission of the Royal Society of New South Wales from *J. Proc. Roy. Soc. N.S.W.*, 1979, **112**, 45-62.

student's) behaviour in class"! Continuous assessment has been with us for longer than some of us may have suspected.

The Liversidge Lecture was bequeathed to us so that we might talk about research of the past and research of the future. What I am going to say will involve aspects of chemistry, biology and physics. I should therefore like to begin by establishing a common vocabulary for what comes later, even though this may involve considerations which many of you will find elementary.

The Jargon of Protein Structures

A protein is a polymer composed of amino acid sub-units which differ from one another only by their side-chains, -R (Table 1). There are 24 types of -R which occur naturally. If the 24 natural amino-acids were the letters of an alphabet, then we could use them to write anything from simple words to complicated sentences. Nature uses the 24 amino acids to assemble a vast variety of molecules representing many levels of complexity.

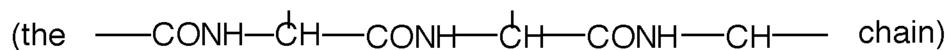


Protein Chemical Structure

TABLE 1
Side-chains of Some Important Amino Acids

Side-chain (-R)	Amino Acid Code	Name	Form at pH7, type
Non-functional: Unlikely to interact with metals			
-H	Gly	Glycine	Non-polar
$-CH_3$	Ala	Alanine	Non-polar, hydrophobic
$-CH_2$- C_6H_5	Phe	Phenylalanine	Non-polar, hydrophobic, aromatic
$-(CH_2)_3$ -ring from C_α to N	Pro	Proline	Non-polar, hydrophobic, cyclic
Functional: Potential metal-binding groups			
$-CH_2$- COO^-	Asp	Aspartate	Acidic, hydrophilic
- $(CH_2)_2$-COO^-	Glu	Glutamate	Acidic, hydrophilic
- CH_2-$CONH_2$	Asn	Asparagine	Uncharged, hydrophilic
$-(CH_2)_2$ -$CONH_2$	Gln	Glutamine	Uncharged, hydrophilic
$-(CH_2)_4$-NH_3^+	Lys	Lysine	Basic, hydrophilic
$-CH_2$-$C_3N_2H_3$	His	Histidine	Basic, pseudo-aromatic
$-CH_2$-C_6H_4OH	Tyr	Tyrosine	Uncharged, polar, aromatic
$-CH_2$-SH	Cys	Cysteine	Uncharged, polar
$-(CH_2)_2$-S-CH_3	Met	Methionine	Non-polar, hydrophobic

The sequence of amino acids in a protein molecule is called the *primary structure*. The local configuration of the 'back-bone'



in a specific region of a protein molecule is called the *secondary structure*. The *tertiary structure* is that arrangement of the primary structure with its elements of secondary structure which minimises the free energy, i.e., it is the molecular structure of the protein.

The best-known element of secondary structure is the α-helix, discovered by Pauling and Corey in the early 1950s. This is a cork-screw arrangement of the protein backbone held together by hydrogen bonds. The side-chains stick out at the sides.

A second important element of secondary structure is one in which the protein backbone is in an extended configuration. Alternate side-chains fall below and above the backbone. If two such backbone segments are adjacent to each other and run in *opposite* directions then there are excellent opportunities for lateral hydrogen-bonding. This arrangement is called 'antiparallel β-structure'. If the adjacent backbone segments run in the same direction, the hydrogen bonding is a little less favourable but still effective. We then have 'parallel β-structure'. Several strands of protein backbone may form what is called a 'β-sheet'. Portions of β-sheet frequently occur at the surfaces of protein molecules. The side-chains of each backbone segment then extend alternatively into the solvent and into the 'interior' of the protein.

Some Aspects of Copper Chemistry

The aspects of copper chemistry which are most relevant to the present story concern some differences between the two most common oxidation states of the metal, +I and +II. Cu(I) complexes ($3d^{10}$) are diamagnetic while Cu(II) complexes ($3d^9$) are paramagnetic. In the 'HSAB' (hard-soft-acid-base) classification, Cu(I) is 'soft'. Cu(II) belongs to the 'intermediate' category between 'soft' and 'hard'.

The most common coordination geometry in Cu(I) complexes is tetrahedral. Linear coordination (coordination number 2) also occurs.

The characteristic coordination geometry of Cu(II) is based on an approximately square and planar arrangement of four ligand atoms. In addition there may be *two, one* or *no* additional ligand atoms lying on an axis perpendicular to the plane of the other four. The axial bonds, if any, are generally longer than the four 'equatorial' bonds. The resulting geometry is elongated octahedral (two axial ligands), square-pyramidal (one axial ligand) or square-planar (no axial ligand).

The crystal structures of many Cu(II) complexes with amino acids and peptides (i.e., protein sub-units) have been studied. Some interesting correlations between ligand-types, coordination numbers and spectroscopic properties have emerged (Freeman, 1967; Billo, 1974). On the other hand, we have never been able to crystallise and determine the structures of amino acid or peptide complexes of Cu(I). Our closest approximation to studying Cu(I) has been a series of structure analyses of the corresponding complexes of another d^{10} metal, Ag(I). The Ag(I) complexes provide examples of linear, trigonal and tetrahedral coordination (Acland and Freeman, 1971; Acland, Flook, Freeman and Scudder, 1972).

Some Properties of Copper-Containing Proteins

Proteins which interact with metal atoms can be divided into two categories:

(i) Proteins which do not contain metal atoms but which are stabilised or potentiated in some way when they combine with metal ions. Such proteins are large organic ligands in equilibrium with metal ions. We shall not consider them further.
(ii) Proteins which have metal atoms as part of their molecular structure: metalloproteins.

The metalloproteins in which the metal is copper have four types of function:

1. Electron transfer.

2. Dioxygen binding.

3. Catalysis.

4. Copper transport.

Functions 1, 2 and 3 all make use of the fact that Cu can exist in two oxidation states which differ by one electron. The number of Cu atoms per protein molecule (or per active subunit where there are several subunits in the molecule) is related to the function of the protein. In electron-transfer Cu-proteins it is 1; in O_2 carriers it is 2; in Cu-enzymes it appears to be 1, 2, 4 or 8; and in ceruloplasmin (a protein whose functions include Cu-transport) it is 6 to 8.

The Cu atoms in most Cu-proteins are distinguished by spectroscopic and redox properties which are bizarre in comparison with the behaviour of Cu(I) and Cu(II) in normal (i.e., low-M.W.) complexes. In the present lecture we shall deal with a 'Type 1' Cu centre. There also exist 'Type 2' and 'Type 3' Cu centres. The molecules of some of the multi-Cu proteins contain Cu centres of all three types. We shall concentrate on the electron-transfer Cu-proteins because they contain only one Cu atom per molecule, and that Cu belongs to 'Type 1'.

The properties which distinguish a 'Type 1' Cu centre are:

(a) An absorption maximum near 600 nm with an absorption coefficient ε_{max} of the order of 5×10^3 $M^{-1}cm^{-1}$. This value is about 100 times higher than for a typical low-M.W. Cu(II) complex.

(b) An abnormally small hyperfine splitting constant $A_{||}$ in the EPR spectrum. The value of $A_{||}$ for a 'Type 1' Cu-protein is characteristically 0.003 - 0.008 cm^{-1}; for a normal Cu(II) complex it is 0.012 - 0.020 cm^{-1}.

(c) A high redox potential, E^0, in the range 0.3 - 0.8 v, compared with the value 0.17 v. for Cu(I)/Cu(II) couple in aqueous solution.

The property which is most likely to be connected with the biological role of 'Type 1' Cu-centres is the redox potential. The intense blue colour and the small hyperfine splitting constant are useful symptoms which help us to diagnose whether we have a 'Type 1' centre or not, but there is no obvious connection between the blue colour of the EPR spectrum and what the proteins are intended to *do*. The function of Cu-proteins with a single 'Type 1' centre - at least in the cases where we think that the function is understood - is to transfer electrons from one redox partner to another. The property which measures the tendency of a molecule to accept or donate electrons is the redox potential.

A number of single-Cu 'Type 1' proteins are shown in Table 2. *Azurin* is a bacterial electron-transfer protein. *Stellacyanin* is obtained from the Japanese lacquer tree. It is an outlier in the Table because its redox potential is only 0.18 v. *Umecyanin* was isolated from horse-radish in Umeå, Sweden. *Plastocyanin* occurs in the leaves of green plants and in some photosynthetic algae. Finally, a *basic blue-green protein* was isolated at about the same time from cucumber seedlings in California and from cucumber peelings in the U.S.S.R.

TABLE 2

Properties of 'Type 1' ('Blue') Copper-Proteins

Protein (Source)	M.W	$\{\lambda_{max}$ (nm)$\}$ $\{\varepsilon_{max}(M^{-1}\ cm^{-1})\}$			$g_{\parallel}$	$g_{\perp}$	$A_{\parallel}$ (cm^{-1})	$A_{\perp}$ (cm^{-1})	E^0 (pH) (mV)
Azurin	14,000	467	666	820	2.26	2.05	0.006	~0	300
(*Ps. aeruginosa*)		27	3500	39					(7)
Stellacyanin	20,000	450	608	850	2.29	2.08,	0.004	0.003,	184
(*Rhus vernicifera*)			4000			2.03		0.006	(7.1)
Umecyanin	14,600	524	606	787	2.32	2.05	0.004	-	283
(Horse radish root)		46	3400	42					(7)
Plastocyanin	10,500	490	597	780	2.23	2.05	0.006	~0	370
(Chloroplasts)		490							(7)
Basic Blue-green	10,100	443	597	750	2.21	2.08,	0.005	0.001,	317
Protein		3500				2.02		0.006	(6.8)
(*Cucumber seedlings*)									

*Includes ~8,000 carbohydrate component

The Crystallographic Study of Plastocyanin: Introduction (Lento) and Scherzo (Allegro)

Even at the present time (July 1978), despite intense efforts in a number of laboratories, no one has synthesised any low-M.W. complex which mimics the unusual combination of properties associated with 'Type 1' Cu centres. It was obvious as long ago as 1970 that the 'Type 1' Cu-proteins have properties which are startlingly different from those of simple model compounds prepared from Cu(II) and amino-acids or peptides. The model compounds are capable of providing much useful, precise and fundamental information, but the structure analysis of a 'Type 1' Cu-protein seemed to offer the only chance of discovering what makes the proteins so different from the models.

The choice of plastocyanin for a structure analysis was not an accident. Firstly plastocyanin has all the properties which are characteristic of the 'Type 1' Cu-proteins. Secondly, at the time when we had to make a choice, plastocyanin had the lowest molecular weight of all the 'Type 1' Cu-proteins, so that the number of atoms which we had to find was minimised. (A marginally lower molecular weight was reported a little later for the blue-green protein from cucumbers.) Thirdly, in comparison with the other 'Type 1' Cu-proteins, plastocyanin has a relatively well-defined biological function; if the photosynthesists are to be believed, then plastocyanin transfers an electron from cytochrome *f* to pigment P700 in one of the steps between Photosystems I and II. Fourthly and finally, a great deal of information about the amino acid sequences of plastocyanins from higher plants and algae was available even in 1971 from the laboratories of Professor Donald Boulter at Durham and Dr. Richard Ambler at Edinburgh. Much of the primary structure (the amino acid sequence) of plastocyanin was known to be more or less invariant. We hypothesised - correctly, as it turned out - that the variations in primary structure would be unimportant in relation to the function of plastocyanin but would be associated with significant differences in crystallisation behaviour.

Our first step in the structure analysis of plastocyanin was to provide ourselves with the experimental material. The isolation and purification of plastocyanin from French bean leaves had been reported by Milne and Wells (1970) at the University of Adelaide. The Hawkesbury Agricultural College was persuaded to sow a crop of French beans. In due course we harvested the leaves. The first preparation of the protein in our laboratory was carried out late in 1971 by an Honours B.Sc. student, Donald Fensom, with help and advice from a number of our friends (see Acknowledgements).

The early experiments did not yield any crystals, but taught us - inorganic chemists and small-molecule crystallographers that we were - some of the facts of life concerning protein chemistry. Three years, several French bean crops, and three Research Assistants later we obtained our first crystals of French bean plastocyanin. There was a brief period of euphoria when everybody admired the beautiful, deep-blue crystals - followed by a long period of gloom: The crystals were long, thin, fragile needles which were unsuitable for diffraction measurements.

Then started the long search for the ideal vegetable: a plant species which is genetically coded to produce plastocyanin molecules with just the right primary structure to give inter-molecular contacts conducive to good crystallisation. Plastocyanins from thirteen plant species were extracted, purified, and subjected to systematic crystallisation experiments over a wide range of conditions. After a while, our agricultural activities were transferred from the Hawkesbury Agricultural College to the University Research Farms at Camden. Silver beet, cauliflower, carrot tops, lettuce, English spinach, cucumber, zucchini, barley, alfalfa ....

This work owed a great deal to the enthusiasm and persistence of two Research Fellows, Dr. John Ramshaw and Dr. M.P. Venkatappa. John Ramshaw never tired of picking up leaves and extracting plastocyanin from them. 'Ven' Venkatappa carried out most of the painstaking crystallisation experiments.

There came a day when the University gardeners pruned the oleander bushes in front of the Chemistry School. The plastocyanin from the oleander leaves yielded beautiful, chunky, stable crystals. Once again success seemed to be within our grasp, until X-ray diffraction photographs revealed that all the crystals were multiple twins and therefore

unsuitable for the structure analysis. We were never able to produce un-twinned crystals of the oleander protein.

The leaves which brought this odyssey to an end came from the poplar trees on the edge of St. Paul's College oval. Poplar plastocyanin yielded large, well-formed crystals (Chapman *et al*., 1977a). The crystals were highly stable in the X-ray beam. They gave X-ray reflections at high angles of reflection corresponding to resolution of 1.6 Å. (Note 1)

There were some uncertainties concerning the antecedents of the poplar trees on St. Paul's oval. It was conceivable that the trees did not belong to the same strain. The crystals used in the structure analysis were therefore grown from a second batch of protein which was extracted from the leaves of a clone (a genetically homogeneous group) of poplar trees in a State forest at Upper Colo.

The details of the structure analysis (Colman *et al.,* 1978) do not concern us here. The preparation of isomorphous heavy-atom derivatives, the recording of data, the massive calculations, and the fitting of a model to the electron-density map at a resolution of 2.7 Å, required a year of intensive work. (Note 2) The persons most directly responsible for the success of the structure analysis are Dr. J. Mitchell Guss and Miss Valerie Norris. It was very helpful, especially in the crucial early stages of the research, that Dr. Peter Colman was working in our laboratory as a Queen Elizabeth II Fellow, so that we could call on his accumulated experience and wisdom.

Generality of the Structural Results

There is a potential criticism which I should like to answer before I describe the results of the structure analysis. It is true that we had indulged in a chemical lottery. We had bet - correctly, as it turned out - that somewhere in the world there was a plastocyanin which would crystallise. Are we entitled to draw any general conclusions about plastocyanin from the structure analysis of the protein from a single plant species chosen in such a haphazard way? The available evidence indicates that we are entitled to do so.

(i) Each plant or other organism which produces plastocyanin is coded so that its plastocyanin has a distinctive amino acid sequence. The plastocyanin molecule has about 100 residues, and there exist 24 amino acids. A lot of combinations are theoretically possible. However, not all the possible combinations occur. There are some positions in the protein chain - for example, position 6 - where the same amino acid is always found. At such a position a particular amino acid has been 'conserved': The primary structure of the protein has changed in response to evolutionary processes, but residue 6 has never been changed successfully from Gly to anything else. Presumably mutants that have a different amino acid at residue 6 do not survive. Similarly, all the plastocyanins that are found in Nature now have a Phe at position 41, an Asn at position 38, and so on. There are, in fact, 28 residues *which never vary* (Boulter *et al.,* 1977). If we omit the algal plastocyanins and concentrate on the plastocyanins from higher plants, then there are 55 residues which are invariant and another 15 where the *type* of residue (hydrophobic, acidic, etc.) is conserved.

The invariance of so much of the plastocyanin molecule supports the hypothesis that the structure of *poplar* plastocyanin represents plastocyanins in general.

(ii) Further evidence comes from a long series of high-field ^{1}H n.m.r. measurements which Dr. Peter Wright in our Department made last year with the collaboration of Dr. John Ramshaw and Miss Valerie Norris (Freeman, Norris, Ramshaw and Wright, 1978). The proteins used for the measurements were the residues of the earlier unsuccessful crystallisation experiments. The availability of relatively large quantities of plastocyanins from about a dozen different plants turned out to be very useful.

The spectra were recorded both for the reduced (Cu(I)) and oxidised (Cu(II)) forms of each protein. In Cu(l)-plastocyanins the metal is diamagnetic so that all the accessible proton signals are recorded. An immediate conclusion from the spectra of the Cu(I)-plastocyanins is that a great deal of the molecular structure must be the same in all of them. There are certainly some significant differences between the spectra: this is to be expected, since no two plastocyanins have precisely the same amino acid sequence and hence precisely the same proton environments. Despite these differences, there are many n.m.r. resonances which persist in the spectra of all the Cu(l)-plastocyanins, showing that the environments of many protons are conserved.

In Cu(Il)-plastocyanins the metal is paramagnetic. This has the result that the proton resonances are broadened differentially depending on the distances of the protons from the paramagnetic centre. The resonances of protons close to the Cu(II) atom are broadened beyond detection. When the (properly scaled) n.m.r. spectrum of a Cu(II)-protein is computer-subtracted from the spectrum of the corresponding Cu(I)-protein, those resonances which are *not* broadened in the spectrum of the oxidised protein disappear. The resonances which are left in the 'difference spectrum' are those which *are* broadened in the spectrum of the oxidised protein, i.e., the resonances of protons close to the Cu(II) atom.

The n.m.r. difference spectra of a series of plastocyanins are almost identical. This shows that, even though there may be variations elsewhere in the plastocyanin molecules, the environments of the Cu atoms in all of them are effectively the same.

'Type 1'Cu Centres: Evidence from 'Sporting' Methods

Even before Dr. Peter Wright's elegant work there had been useful ^{1}H n.m.r. experiments on plastocyanin and other 'Type 1' Cu-proteins. Following a visit to our Department by Dr. H.A.O. Hill from Oxford in 1973, Don Fensom (by now a post-graduate student) spent six weeks in Dr. Hill's laboratory during 1974. Working with Dr. Hill and his colleagues, Don Fensom recorded the high-field n.m.r. spectra of Cu(I)- and Cu(II)plastocyanins at a series of pH's. The chemical shifts of two of the proton resonances in the spectrum of the Cu(l)-protein had pH dependences which suggested that the protons belonged to the imidazole rings of two histidine residues. These particular resonances were among the first to be broadened beyond detectability when the protein was oxidised to the Cu(II) form. More extensive experiments of the same kind were reported by Dr. John Markley at Purdue. This was the first evidence that two imidazole groups are close to the Cu atom (Beattie *et al.*, 1975; Markley *et al.,* 1975).

Applications of other spectroscopic techniques also led to interesting conclusions concerning the Cu site in plastocyanin. In one ingenious experiment, Dr. Harry B. Gray and co-workers at the California Institute of Technology recorded and compared the ESCA spectra of Cu(II)-plastocyanin, Co(Il)-substituted plastocyanin and metal-free ('apo') plastocyanin (Solomon *et al.*, 1975). A major peak in the spectrum of apo-plastocyanin

was identified as arising from the 2p electrons of three sulfur atoms. This observation was consistent with the presence of one cysteine and two methionine residues in the amino acid sequence. The sulfur 2p peak was reduced by about one-third in the Cu(II)- and Co(II)-plastocyanin spectra; simultaneously a satellite peak appeared, with an integrated area about one-third as large as the original peak. It was concluded that one of the three sulfur atoms in the molecule - probably the thiol sulfur of the cysteine side-chain - is bonded to the Cu atom. It is unfortunate that a subsequent reassessment revealed a flaw in the ESCA experiment and that the original observations have a different explanation.

Nevertheless, at the time when the ESCA results were reported they were consistent with two other pieces of evidence for the involvement of a cysteine thiol group in Cu-binding. Firstly, it had been shown earlier that apo-plastocyanin recombines easily with Cu(II) but that the recombination is prevented by thiol-specific mercurial reagents. Secondly, a thiol-Cu bond could account for the very high extinction coefficient of the absorption band near 600 nm. This band could then be attributed to S -----> Cu charge transfer.

The existence of a strong CT chromophore led logically to the recording of resonance Raman spectra for a number of 'Type 1' Cu-proteins. Similar spectra were reported by two laboratories. Dr. T.G. Spiro and co-workers at Princeton interpreted the spectra on the basis of a trigonal-bipyramidal coordination geometry involving four nitrogen or oxygen donor atoms in addition to a cysteine sulfur (Mistowski *et al.*, 1975). Dr. M.N. Young's group at Ottawa obtained an explanation (equally plausible, so far as I can judge) by using a tetrahedral Cu geometry with only three donor atoms in addition to the sulfur (Siiman, Young and Carey, 1976).

The cumulative result of all these experiments was that three of the Cu-binding groups in plastocyanin were tentatively identified as two histidine imidazole groups and a cysteine thiol group. Two further experiments led to the suggestion that the Cu atom forms a fourth bond to the de-protonated nitrogen atom of an amide group in the protein backbone. One of these experiments involved the recording of the infrared spectra of Cu(I)-, Cu(II)-, Co(II)- and apo-plastocyanin (Hare, Solomon and Gray, 1976). Supporting evidence for Cu-amide binding came from the ^{13}C n.m.r. difference spectrum for the Cu(I) and Cu(II) forms of a related protein, azurin (Ugurbil *et al.*, 1977). A 1976 'state of the art' diagram of the Cu site in plastocyanin showed an approximately tetrahedral arrangement of a cysteine sulfur, two imidazole nitrogens and amide nitrogen (Solomon, Hare and Gray, 1976).

The Structure of Plastocyanin

On the afternoon of Sunday, September 2, 1977 Mitchell Guss and Valerie Norris finished the job of fitting the first complete model of plastocyanin to the electron-density map at 2.7 Å resolution. The results were telephoned to me in London (where it was still Sunday morning). I spent the rest of the day building a wire model according to the telephoned data. The structure was unveiled at the International Congress on Photosynthesis in Reading three days later (Colman *et al.,* 1977c). It was enormously gratifying that the audience included two friends who had in different ways shared the frustrations of the long period when we had no crystals: Dr. Allen Hill, who drove from Oxford, and Dr. John Ramshaw, who flew in from Harvard.

I have said that the structure, in the form in which it was reported at Reading and subsequently published, was solved "at 2.7 Å resolution". It is important for chemists and biologists to appreciate what "2.7 Å resolution" means. It means that the structure is viewed as though we used a magnifying glass which can resolve objects only when they are separated by distances larger than about 2.7 Å. For example, the carbon atom and oxygen atom of a C=O bond, 1.3 Å apart, will not be resolved. On the other hand, the resolution is sufficient to make a distinction between the $-CH_3$ side chain of an Ala and the $-CH_2C_6H_5$ side chain of a Phe, and in many cases there will be bumps in the electron-density at places where the amide C=O groups protrude from the protein backbone.

In the case of plastocyanin, the electron density map at 2.7 Å resolution told us what the chemical ligands of the Cu atom are, but did not yield meaningful values for the bond-lengths and bond-angles.

The poplar plastocyanin molecule comprises about 800 atoms. The structure is most easily understood if it is initially represented in a highly idealised way. Let us delete all the side-chains, leaving only the protein backbone; let us then delete the amide groups, leaving only the C atoms; let us smoothe out the kinks and irregularities in the chain; and let us finally adapt what is left to the surface of a flattened barrel, which happens to be the regular object that it most closely resembles (Fig. 1).

The molecule is seen to consist of eight strands of protein backbone with bends between the strands. These eight strands form the walls of what we have just described as a barrel. Starting at the $^+NH_3$-end of the protein chain, strand 1 goes up along the front of the barrel. Strand 2 goes down. Strand 3 goes up and over the top, and strand 4 comes down on the other side. Strand 5 (up, at the back of the barrel) is irregular. Strand 6 (down, at the side) leads to a loop under the barrel. Strand 7 (up) ends in the double loop, after which strand 8 continues down to the $-COO^-$ terminus.

The seven strands other than strand 5 have considerable β character. At 2.7 Å resolution the details of the hydrogen bonding are not yet defined with precision, but many of the side-chains are clearly visible in the electron-density map. The evidence for β character is that alternate side-chains appear on opposite sides of the backbone. Strand 5

FIGURE 1. Highly schematic representation of the plastocyanin molecule. The $^{+}NH_3$-terminal residue is denoted by N, and the COO^{-} -terminal residue by C. The Cu atom is represented by the black ball.

is quite irregular. It appears to have no β character, and hangs a little outside the rest of the barrel.

The Cu atom lies near one end of the barrel, slightly below the boundary defined by the loop between strands 3 and 4 and the double loop between strands 7 and 8.

Figure 2 is a diagram in which the protein chain is drawn by connecting the actual - not idealised - positions of the C_α atoms. Individual amino acids are represented by their C_α atoms. Some of them are labelled to indicate their positions in the sequence. The Cu atom is coordinated by the side-chains of residues 37, 84, 87 and 92. The coordination geometry of the Cu atom has been drawn to resemble a tetrahedron, but all that we can say at 2.7 Å resolution is that it is irregular: it is less unlike a tetrahedron than it is unlike a square!

The four donor atoms bounded to the Cu atom are

- the δ-nitrogen of the imidazole group of His 37,
- the δ-nitrogen of the imidazole group of His 87,
- the thiol sulfur of Cys 84, and
- the thioether sulfur of Met 92.

The participation of the two His nitrogens and the Cys sulfur is in accordance with the predictions from spectroscopic and chemical observations. The ligand group which surprised almost everybody was Met 92. Indeed, for some days after its discovery as a ligand, the thioether group of Met 92 was a distinct embarrassment. It is worth making a small digression to explain why this was so.

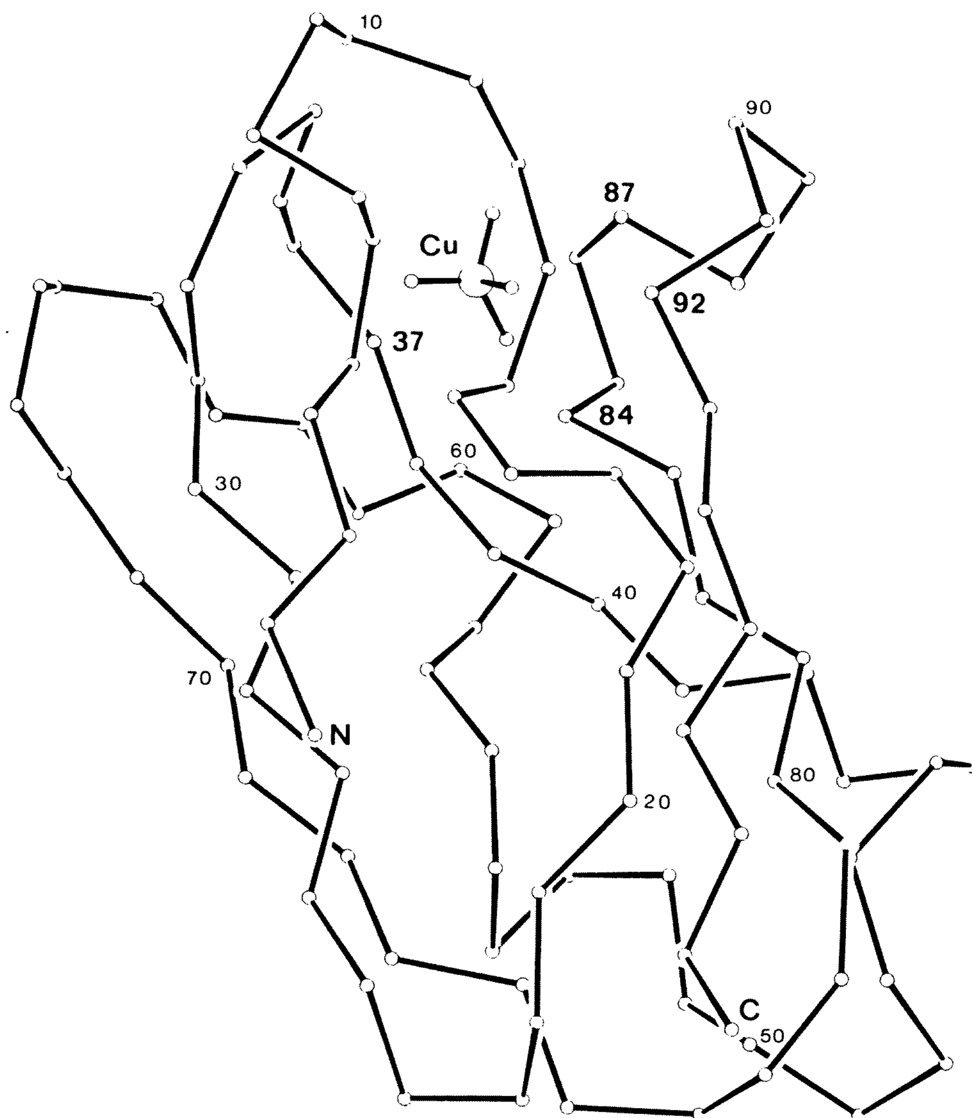


FIGURE 2. The plastocyanin molecule, drawn by linking the C_{α} atoms of the 99 amino acid residues. Small numerals identify every tenth residue. Large numerals identify the four Cu-binding residues. The Cu atom is shown with the four donor atoms to which it is bonded. (This drawing was made in March, 1979 using the atomic coordinates obtained by refinement at 1.6 Å resolution.)

Among the plastocyanins which had been sequenced at Durham was the plastocyanin of a weed called 'dock'. It differed from the other plastocyanins which had been sequenced

in having a Leu instead of a Met at residue 92. This result had two possible implications. If, as seemed likely, *all* plastocyanins used the same functional groups to coordinate the Cu atom, then Met 92 could not coordinate the Cu atom in any of them since it was absent in dock. Alternatively, if Met 92 was a Cu-binding residue in one or more plastocyanins, then there was at least one species which had to coordinate the Cu atom in some other way; the generality of the structure analysis of poplar plastocyanin would be lost.

On the Sunday when Mitchell Guss transmitted the structure to me by phone from Sydney to London, he also called John Ramshaw in Cambridge, Mass. Ramshaw took the next convenient flight to England, so as to be present when the structure was announced at the International Congress on Photosynthesis. After arriving in England he called at the University of Durham and, with Professor Boulter's permission, checked the original laboratory records of the amino acid sequence determination for dock plastocyanin. It turned out that there had been an error in the interpretation of the sequence. Dock plastocyanin, like all the other plastocyanins of which we are aware, has a Met at residue 92.

The Usefulness of 'Model' Compounds

To what extent could a distorted tetrahedral coordination with two nitrogen and two sulphur donors have been predicted from properties of simpler complexes? In retrospect, all the main features of the Cu coordination in plastocyanin were predictable and had been predicted. What caused difficulties was not a lack of predictions, but an excess.

(i) Several authors had pointed out that there is a possible analogy between the intense visible absorption bands of 'blue' Cu-proteins and the S---->Cu charge-transfer bonds of certain Cu(II)thiol complexes. For example, the structure of a Cu(I), Cu(II) mixed-valence complex of a β,β-dimethyl-D-cysteine (D-penicillamine) was solved by Dr. Paul Birker in our laboratory in 1976 (Birker and Freeman, 1977). Each of the six Cu(II) atoms in the complex is coordinated by two thiolate sulphur atoms and two amino groups. The extinction coefficient at 518 nm is about 5000 $M^{-1}cm^{-1}$ per Cu(II) atom.

(ii) The abnormal EPR spectra of 'Type 1' Cu-proteins are another property for which a plausible explanation could be (and had been) found from low-M.W. complexes. Gould and Ehrenberg (1968) in Stockholm examined the effect of irradiating tetrakis-(acetonitrile)- copper(I) perchlorate with γ-rays. The complex is colourless. The Cu(I) atoms are diamagnetic, i.e., they are EPR-inactive. The coordination geometry of the Cu(I) atoms has been shown to be tetrahedral by structure analysis. γ-Irradiation knocked an electron out of some of the Cu(I) atoms, thereby creating Cu(II) centres in tetrahedral environments. The Cu(II) centres were paramagnetic and EPR-active. In the EPR spectrum of the irradiated complex, the hyperfine splitting constant $A_{||}$ was 0.008 cm^{-1} - at the upper limit for 'Type 1' Cu-proteins, and well below the limit for normal Cu (II) complexes.

(iii) A number of reasonable explanations of the high reduction potentials of 'Type 1' Cu-proteins were embodied in a series of E^0 measurements for the Cu^+/Cu^{2+} couple in aqueous solutions containing various organic ligands (James and Williams, 1961). In the absence of ligands the value of E^0 is 0.167 v. Run-of-the-mill amino acids cause a drop to negative values since they stabilise Cu(II) by forming excellent chelate complexes. In the presence of imidazole the value of E^0 is increased to 0.35 v., close to the values for some

of the 'Type 1' Cu-proteins. A high redox potential is also found in the presence of 2,9-dimethyl-1,10-phenanthroline. This ligand and Cu(II) form a strong 2:1 complex in which the normal square-planar coordination geometry of Cu(II) is distorted towards a tetrahedron, due to steric hindrance between the methyl groups of the two ligand molecules. The tetrahedral distortion de-stabilises Cu(II) with respect to Cu(I). Yet another ligand which causes a large increase in E^0, is ethylenebisthioglycollic acid, which has two thioether groups. In other words, long-established data on model compounds included evidence that coordinated imidazole groups or coordinated thioether groups or a tetrahedral distortion (or presumably a combination of any of these factors) are consistent with the high redox potentials of the 'Type 1' Cu-proteins.

(iv) A series of Cu(II) complexes with cyclic thioether ligands had been reported by Dr. David Rorabacher at Wayne State University in 1975 (Jones *et al.*, 1975). In these complexes the number of sulfur donors per Cu(II) ranges from 1 to 4. All the complexes have high redox potentials in the same range as the 'Type 1' Cu-proteins, and exhibit intense charge-transfer absorptions (though these occur at slightly shorter wave-lengths than in the case of the proteins). The crystal structure analysis of one complex had shown that the Cu(II) atom has a square-planar geometry. This work had drawn attention to thioether coordination and had queried the importance of thiol coordination and tetrahedral distortion - as a possible cause of the spectroscopic and redox properties of the 'Type 1' Cu-proteins.

In summary, model compounds would have enabled us to understand the 'Type 1' Cu-proteins if we had known which models we should take seriously.

Elegance in Molecular Design

The title of this lecture is 'Elegance in Molecular Design'. Nature 'designs' things by trying out random variations and getting rid of the variants which don't work. In the case of plastocyanin, the result is a splendid example of structure idealised for function. Consider the Cu-binding site (Fig. 3). Above the Cu (or, if you prefer, to the North) lies the imidazole ring of His 87. This is all that separates the Cu atom from the world outside its molecule. Below the Cu (or towards the South, in the direction of the body of the molecule) are the imidazole group of His 37, the thiolate sulphur of Cys 84, and the thioether sulfur of Met 92. How is this combination of ligands related to the function of the protein?

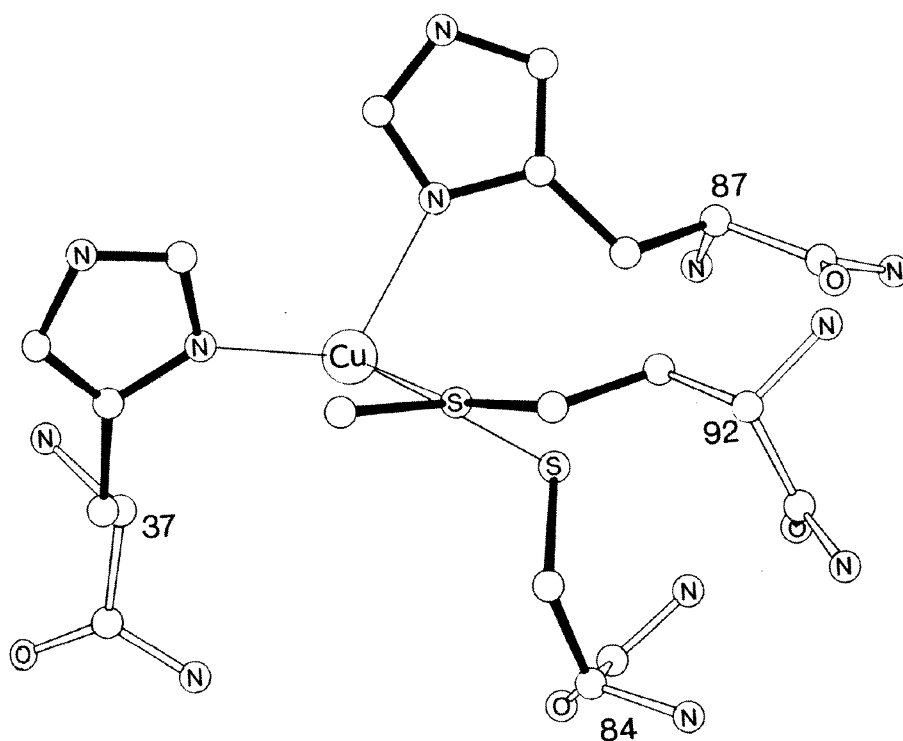


FIGURE 3. Detailed view of the Cu atom and the four Cu-atoms binding groups in plastocyanin. Black bonds are in the amino acid side-chains. Bonds not drawn in black are parts of the protein back-bone.

The function of plastocyanin is to transfer electrons. Electron-transfer clearly depends on reversible changes between the two oxidation states of the Cu atom. The first achievement of the molecular design process is that the Cu atom has been given two sulfur donor atoms which are ideal for Cu(I) and acceptable for Cu(II), and two rather basic imidazole nitrogen atoms which are ideal for Cu(II) and acceptable for Cu(I). Changes from Cu(II) to Cu(I) and from Cu(I) to Cu(II) can almost certainly be accommodated by minor local conformational changes, without the breaking and making of metal-ligand bonds. The coordination sphere excellently suits a Cu atom whose purpose in life is to undergo reversible changes in oxidation state.

Reactions in which a metal changes its oxidation state without the rupture or formation of metal-ligand bonds are called "outer-sphere" electron-transfer reactions. Outer-sphere electron-transfer means that two complexes come into sufficiently close contact to establish significant orbital overlap, and that an electron is delocalised from the metal centre of one complex to the metal centre of the other. The 'northern'/edge of the imidazole ring of His 87 would be a useful point of contact between plastocyanin and an outer-sphere redox partner, because the imidazole ring provides a conjugated pathway to and from the Cu atom.

A feature of the Cu site which may make some inorganic chemists feel uneasy is the distorted coordination geometry. We shall not know the precise geometry until the structure is refined, but even at 2.7 Å reduction we can see that the coordination of the Cu atom is grossly distorted from the square-planar geometry preferred by Cu(II) towards the tetrahedral geometry preferred by Cu(I). As we have already seen in the case of low-M.W. complexes, a distortion from square-planar coordination geometry increases the redox potential of the Cu(I)/Cu(II) couple, i.e., increases the tendency of Cu(II) to accept an

electron. This is just what is needed in plastocyanin, which has to provide the electron-transfer link between two high-potential components of the photosynthetic chain. By the standards of small-molecule geometry, the coordination of the Cu atom is strained; but the strain is built into the Cu site as part of Nature's molecular design.

What is the origin of the strain energy? Protein molecules have more sources of free energy than the small molecules with which inorganic coordination chemists usually deal. A protein molecule has many hydrogen bonds, interactions between side-chains bearing opposite charges, and contacts between hydrophobic groups. Free energy saved in one part of the molecule can be used to pay for strain which is needed in another part. This is the entatic state hypothesis (Vallee and Williams, 1968), and the Cu-binding site in plastocyanin provides a good example of it.

Invariant Amino Acid Residues

Special interest is attached to those 55 amino acid residues which are conserved in the higher plant plastocyanins (Fig. 4). Presumably each one of them is conserved for good reasons. It is possible to test this hypothesis by making a close inspection of the model of the plastocyanin molecule. Let us, however, keep in mind that our present model is fitted to the electron-density at a resolution of only 2.7 Å.

5 10 15 20 25
Ile-Asp-Val-Leu-Leu-Gly-Ala-Asp-Asp-Gly-Ser-Leu-Ala-Phe-Val-Pro-Ser-Glu-Phe-Ser-Ile-Ser-Pro-Gly-Glu

30 35 40 45 50
Lys-Ile-Val-Phe-Lys-Asn-Asn-Ala-Gly-Phe-Pro-His-Asn-Ile-Val-Phe-Asp-Glu-Asp-Ser-Ile-Pro-Ser-Gly-Val

55 60 65 70 75
Asp-Ala-Ser-Lys-Ile-Ser-Met-Ser-Glx-Glx-Asx-Leu-Leu-Asn-Ala-Lys-Gly-Glu-Thr-Phe-Glu-Val-Ala-Leu-Ser

80 85 90 95
Asn-Lys-Gly-Glu-Tyr-Ser-Phe-Tyr-Cys-Ser-Pro-His-Gln-Gly-Ala-Gly-Met-Val-Gly-Lys-Val-Thr-Val-Asn

POPLAR (*Populus nigra* var. *italica*) PLASTOCYANIN SEQUENCE

FIGURE 4. The amino acid sequence of poplar plastocyanin (R. Ambler, personal communication), as used in the analysis of the structure. Some minor ambiguities remain to be resolved (Glx = Glu or Gln Asx = Asp or Asn). The residues in a few short segments have been ordered by analogy with other plastocyanin sequences.
Doubly underlined residues are conserved in all the higher plant plastocyanins for which we have sequence data. Singly underlined residues are conserved so far as the type of side-chain (acidic, non-polar, aromatic, etc.) is concerned. See, however, Table 3 for notes about residues 24, 45 and 95.

(i) Invariant Gly residues.

Nine Gly residues are totally conserved. Another is conserved with only one known exception. Gly residues are subject to fewer geometrical constraints than amino acids with bulkier side-chains. For this reason they facilitate the formation of bends or turns in the protein backbone. In plastocyanin, nine of the ten conserved Gly residues are found in seven of the bends where the protein chain changes direction (Table 3(a)).

(ii) Invariant Pro residues.

The geometry of Pro is constrained in a special way by the cyclic side chain, so that Pro residues frequently occur at the ends of secondary structures such as α-helices and β-sheets since they cannot be accommodated in the middle. All four conserved Pro residues in plastocyanin occur at places where the protein backbone undergoes a turn [Table 3(b)]. Two of them, Pro 36 and Pro 86, appear to have a special function which will be mentioned shortly.

(iii) Conserved non-polar residues.

In addition to the four Pro residues mentioned above, there are 22 other conserved non-polar residues (Ala, Ile, Leu, Met, Phe and Val). They include Met 92, one of the four residues to which the Cu is attached. There are another five positions in the sequence which are always occupied by one of the non-polar residues Ala, Ile or Val. In other words, 31 residues are invariant as to hydrophobic character, and 26 of them are totally invariant.

In seventeen cases, the hydrophobic sidechains of these residues point into the interior of the molecule. The molecule is not only a flattened barrel. It is a flattened *oil* barrel. This result is not particularly remarkable. Many other protein molecules are known to have predominantly hydrophobic interiors.

On the other hand, a number of the hydrophobic residues in plastocyanin are associated with a distinctive patch near the Cu site. The Cu atom is located at the bottom of a shallow crater. The rim of this crater is lined by four hydrophobic residues and four Gly's. Three more hydrophobic residues lie on the sides of the crater. All these residues are conserved. They include two of the invariant prolines, Pro 36 and Pro 86 [Table 3(c)]. In this region of the molecule there are no charged side-chains at all.

The function of the hydrophobic patch near the Cu site may be *either* to orient the plastocyanin molecule in or on the thylakoid membrane, or to make the business end of the molecule recognisable by one or other of plastocyanin's redox partners. The need for a recognition patch is suggested by the high degree of specificity which has been reported to exist between plastocyanin and its natural redox partner, cytochrome *f*. Electron transfer from cytochrome *f* to plastocyanin has been found to be 30 times faster than from any other cytochrome; and electron transfer from cytochromes is generally 1000 times faster than from artificial reducing agents such as inorganic complexes.

(iv) Invariant polar uncharged residues.

There are ten conserved polar uncharged residues (Asn, Gln, Ser, Thr, Tyr) (Table 3(d)). In only one case can we see a clear reason why the residue is conserved: Asn 38 has the C=O of its side-chain amide group hydrogen-bonded to the nitrogen of the backbone amide group of Ser 85. Asn 38 is adjacent to the Cu-binding residue His 37, while Ser 85 is adjacent to the Cu-binding residue Cys 84. The side-chain of Asn 38 thus stabilises the configuration of the Cu-site by linking two strands of the protein backbone near two of the Cu ligands.

Six of the remaining conserved polar uncharged residues appear to be in contact with the solvent. Two of them, Tyr 83 and Gln 88, will be mentioned again later in connection

with the 'negative patch' on the plastocyanin molecule. The side-chain of Tyr 80 points to the interior of the molecule.

(v) Invariant basic residues.

The amino-acids which have ionisable side-chains are distributed unequally throughout the plastocyanin molecule. In poplar-plastocyanin there are only six basic residues: two His and four Lys. Five of these residues are conserved in higher plant plastocyanins (Table 3(e)). They are (i) His 37 and His 83, which bind the Cu atom through their imidazole groups, and (ii) Lys 30, Lys 54 and Lys 95, which point from the surface of the molecule into the solvent.

TABLE 3

Some Amino Acid Residues which are Invariant in Higher Plant Plastocyanins

Residue	Location of Residue or Function of Side-chain
(a) Invariant Glycines	
Gly 6, Gly 10	Bend between strands 1 and 2
Gly 24*	Bend between strands 2 and 3
Gly 34	Bend between strands 3 and 4
Gly 47	Bend between strands 4 and 5
Gly 67	Bend between strands 5 and 6
Gly 78	Bend between strands 6 and 7
Gly 89, Gly 91	Bend between strands 7 and 8
Gly 94	β-structure
(* Residue 24 in beetroot plastocyanin is Ser.)	
(b) Invariant Prolines	
Pro 16	Bend in strand 2
Pro 36	Beginning of strand 4
Pro 47	End of strand 4
Pro 86	Loop between strands 7 and 8
(c) 'Hydrophobic Patch'	
Gly 10	Rim of Cu pocket
Leu 12	Rim of Cu pocket
Ala/Val 13	Rim of Cu pocket
Phe 14	Side of Cu pocket
Gly 34	Rim of Cu pocket
Pro 36	Side of Cu pocket
Leu 62	Side of Cu pocket
Pro 86	Rim of Cu pocket
Gly 89	Rim of Cu pocket
Ala 90	Rim of Cu pocket
Gly 91	Rim of Cu pocket

TABLE 3 (continued)

(d) Invariant Polar Uncharged Residues

Residue	Function
Asn 31	Function?
Asn 32	Contact with solvent
Asn 38	Link from strand 3 to Ser 85 in strand 7
Ser 56	Function?
Asn 64	Contact with solvent
Tyr 80	Interior of molecule
Tyr 83	Contact with solvent near 'negative patch'
Gln 88	" " " " " "
Thr 97	Contact with solvent
Asn/Gln 99	Contact with solvent, -COO^- terminus

(e) Invariant Basic Residues

Residue	Function
Lys 30	Contact with solvent
His 37	Cu-binding group
Lys 54	Contact with solvent
His 87	Cu-binding group
Lys 95*	Contact with solvent

(*Residue 95 in broad bean plastocyanin is Gln)

(f) Invariant Acidic Residues

Residue	Function
Asp/Glu 2	Function?
Glu/Asp 25	Contact with solvent
Asp 42	Contact with solvent, side-chains form part of 'negative patch'
Glu 43	" " " " " " " " " "
Asp 44	" " " " " " " " " "
Glu 45*	" " " " " " " " " "
Asp S1	Contact with solvent
Glu 59	Contact with solvent, side-chains form part of 'negative patch'
Asp/Glu 61	" " " " " " " " " "
Glu 68	Function ?
Cys 84	Cu-binding group

(* Residue 45 is reported to be Ser in poplar plastocyanin)

(vi) Invariant acidic residues

The amino acids conventionally described as acidic are Asp and Glu. In the case of plastocyanin we may add Cys (which is usually included in the 'polar uncharged' category), since the thiolate group of Cys 84 functions as a Lewis acid by binding the Cu atom. There are then eleven positions in the plastocyanin sequence where an acidic residue is always found: ten Asp/Glu and one Cys (Table 3(f)). Some doubt is still attached to residue 45 which is conserved as Glu, except in poplar plastocyanin where it is reported to be Ser.

In eight of the ten conserved residues with -COO^- groups, the side-chains are directed into the solvent. This is as expected. What is perhaps unexpected is that the acidic residues are significantly more on one side of the molecule than on the other. Six of the conserved Asp/Glu residues, plus Gln 88 (which is also conserved - see above) form a distinctive, elongated, negative patch. We have asked ourselves whether such a distinctive structural feature has a functional significance.

Plastocyanin should have *two* electron-transfer pathways - one to get an electron in, and the other to get it out. Since two different redox partners are involved in these processes it is unlikely that they make contact with the plastocyanin molecule at the same place. One possible contact point, namely the exposed edge of the His 87 imidazole ring, has already been discussed. The patch of conserved negative residues on the side of the plastocyanin molecule may be another.

This idea was first suggested to us by Dr. Peter Wright. A conserved tyrosine residue, Tyr 83, has its aromatic side-chain in contact with the solvent at about the centre of the negative patch. From that point to the Cu atom there appears to be a straight channel lined by the methylene group of Gly 94, the phenyl rings of Phe 14 and Phe 82, and the aliphatic side-chain of Val 93. Such a channel filled with a medium having a low dielectric constant is a requirement for electron-transfer by "quantum mechanical tunnelling". The existence of a hydrophobic channel does not prove that quantum mechanical tunnelling occurs. It merely satisfied one of the prerequisites.

It may be objected that, since most of the interior of the plastocyanin molecule is filled with hydrophobic side-chains, there are *many* paths through regions of low dielectric constant. What makes the path from the negative patch to the Cu so special is that the four residues which I have mentioned are *invariant*. We note that they include Gly 94 which, unlike the other nine invariant Gly's in plastocyanin, is not found in a bend of the protein chain. The need for a hydrophobic channel from the Cu atom to a negatively charged recognition patch on the surface of the molecule would provide a reason why these four particular residues, including Gly 94, are conserved.

Summary

The tentative conclusions which we draw from the poplar plastocyanin structure at 2.7 Å resolution are as follows: The plastocyanin molecule has evolved to produce a Cu centre which can accept and give up electrons without changes in coordination and with minimal changes in geometry. The Cu centre is not in direct contact with the world outside the protein molecule. There are at least two pathways along which electrons may be transferred to and from the Cu centre by mechanisms which are compatible with the contemporary folk-lore of inorganic electron transfer reactions. Both these pathways terminate in recognisable patches - at present, the *only* recognisable patches - on the surface of the molecule. (Note 3).

Future Prospects

A Liversidge Lecture should indicate some directions in which research might proceed from here.

(i) We hope to continue our calculations, incorporating higher resolution data which have already been recorded. This will make the structure analysis more precise. We shall then be able to attach meaningful values to the metal-ligand bond-lengths and bond-angles; we shall be able to study the details of the β-structure (because we shall know just where the hydrogen bonds are); and we shall be able to describe the intermolecular contacts in the crystal, which may lead to an understanding why poplar plastocyanin crystallised so nicely but thirteen other plastocyanins did not.

(ii) If the 'refinement' calculations in (i) lead to a sufficiently precise knowledge of the structure, then it will probably be possible to detect how the protein responds structurally to various chemical changes. For example, we already know that Cu(I)-plastocyanin is isomorphous with Cu(II)-plastocyanin. The measurement of X-ray diffraction data for Cu(I)-plastocyanin to high resolution should therefore enable us to calculate the changes in electron-density (i.e. changes in structure or conformation) which accompany electron-transfer. Similarly it may be possible to determine whether the protein conformation is affected by pH; and it may be possible to locate the binding sites of certain inorganic redox reagents which are known to associate strongly with the protein before electron transfer takes place.

(iii) It would be interesting to study the structure of stellacyanin, a glycoprotein with a 'Type 1' Cu centre. This protein lacks Met residues completely so that at least one of the ligand groups at the Cu centre cannot be the same as in plastocyanin. We have not yet succeeded in crystallising stellacyanin, but we do have crystals of a new basic green-blue protein from cucumber seedlings (Colman *et al.*, 1977*b*). Both stellacyanin and the cucumber protein have spectroscopic properties which are significantly different from those of plastocyanin. An understanding of the structural reasons for the spectroscopic differences between plastocyanin, stellacyanin and the basic cucumber protein may help to improve the predictive values of the various spectroscopic techniques when applied to metalloproteins.

(iv) In the long term it will be necessary for someone to study the structures of Cu-proteins with 'Type 2' and 'Type 3' Cu centres. The enzyme laccase is particularly attractive from this point of view since it contains a 'Type 1' centre, a 'Type 2' centre and a (double Cu) 'Type 3' centre. At this time no one has been able to induce laccase to crystallise in a form which diffracts X-rays.

It is a chastening epilogue to this lecture that the nucleation of crystals - a chance event which we cannot yet understand, predict or control - still stands between us and the investigation of Nature's molecular designs.

Acknowledgements

I am pleased and grateful that this Liversidge Lecture gives me the opportunity to acknowledge the contributions and assistance of many colleagues.

First and foremost I thank those who have worked on the plastocyanin project; *Donald* (now *Dr. D. J.) Fensom* (Honours student, 1971), who isolated our first specimen of French bean plastocyanin and later - as a Ph.D. student - recorded high-field n.m.r. spectra at Oxford and kinetic data at the California Institute of Technology; *Tad Bohdanowicz* (Research Assistant, 1972); *Elizabeth (now Dr. E.J.) Woodcock* (Research Assistant, 1973); *Dr. Graeme Chapman* (Professional Officer, 1974-5), who set up our

protein chemistry laboratory in its present form; *Dr. John Ramshaw* (Research Fellow and Professional Officer, August 1974 - February 1977), who brought to the research a superb feeling for Cu-protein biochemistry, and maintained a passion for purifying plastocyanin from all sorts of leaves until poplar plastocyanin finally yielded crystals; *Dr. M.P. Venkatappa* (Research Fellow, 1975-7), whose patience and attention to detail were essential for the long series of systematic crystallisation experiments and the production of isomorphous heavy-atom derivatives; *Alan Watson* (Research Assistant, 1975); *Dr. Peter Colman* (Queen Elizabeth II Fellow and later N.H. & M.R.C. Research Fellow, 1975-8), a distinguished protein crystallographer working on his own project in our laboratory, who gave us a great deal of his know-how and active collaboration; *Dr. Mitchell Guss* (Professional Officer, 1975 --), who deserves most of the credit for the structure analysis; *Valerie Norris* (Research Assistant (1976 --); and *Dr. Mitsuo Murata* (Professional Officer, 1976--). The three last-named colleagues are still continuing the work.

Dr. James K. Beattie, Dr. Peter Wright and *Dr. Carolyn Wright-Mountford* are departmental colleagues who were particularly helpful when we came to evaluate and interpret our results. *Dr. Julian Wells* at the University of Adelaide, who isolated and characterised French bean plastocyanin in 1971, gave us helpful telephoned advice during our early protein isolation experiments. *Dr. Bill O'Sullivan* (now *Professor W.J. O'Sullivan* of the School of Biological Sciences at the University of New South Wales) was the person who first pushed Milne and Wells' plastocyanin paper under my nose. *Mr. J. Sumeghy* at the Hawkesbury Agricultural College showed us how to grow French beans in 1971 and 1972. *Dr. Frank Crofts*, Director of the Sydney University Research Farms at Camden, made fields, advice and labour available for the production of crops of numerous vegetables from 1973 to 1976. Experimental facilities at the C.S.I.R.O. Division of Food Research, North Ryde, were made available through the courtesy of *Mr. M.V. Tracey*, Chief of the Division, in 1971-2 when we had not yet established our own protein laboratory. For some years beginning in 1971 we also benefited from having *Mr. Malcolm Smith* at the C.S.I.R.O. Division of Food Research to hold our inexperienced inorganic hands while we extracted and purified our protein specimens.

Much of our laboratory's folk-lore concerning protein crystallisation dates back to a very useful one-month visit by *Mr. Larry Sieker* in 1975. For the purpose of that visit he was helpfully released from his normal duties in *Professor Lyle Jensen's* laboratory at the University of Washington, Seattle. *Professor Jan Drenth* at Groningen and *Professor Robert Huber* at Munich supplied us with computer programs which were subsequently adapted to our needs and computer configuration and used in the structure analysis. *Dr. Richard Ambler* at Edinburgh determined the amino acid sequence of poplar plastocyanin. Two of my friends in the Cu-protein business, *Dr. Allen Hill* at Oxford and *Professor Harry Gray* at the California Institute of Technology, provided challenge, stimulation and encouragement over a long period.

An unforgettable example of fraternal generosity was provided by *Professor Hans Jansonius* at the Biozentrum in Basel, Switzerland. I was visiting Basel in 1976 when Mitchell Guss cabled me that poplar plastocyanin had been successfully crystallized at last. Due to an ambiguity in the cable I thought that we had only ten crystals and that the next Australian poplar season was a year away. Basel, however, was full of leafy poplar trees. Although Professor Jansonius was busy preparing an important lecture for an international meeting, he immediately mobilised his research group, obtained the official blessing of the Canton of Basel, and stripped the cantonal poplars of their leaves. In the end we did not take advantage of this supply of raw material: the apparent lack of poplar

leaves in Sydney was shown to be the result of a misunderstanding, and the Basel poplars turned out to belong to a different species than those in Sydney.

Finally, it is a pleasure to record that the study of the crystal structures of metalloproteins at Sydney University has been supported by the Australian Research Grants Committee, by the University of Sydney through the University Research Grant and a General Development Grant, and in 1978 by a donation from Esso Ltd.

Notes

Note 1. The diffraction process is governed by the Bragg equation, $n\lambda = 2d \sin\theta$. The larger the angles θ at which reflections can be recorded, the smaller are the spacings d which can be resolved.

Note 2. X-ray data sets were recorded for the native protein and for three heavy-atom derivatives which were isomorphous with the protein. The derivatives were obtained by soaking protein crystals in solutions of mercuric acetate, uranyl acetate for 23 h, and uranyl nitrate for 10 d, respectively. The X-ray data were initally recorded to a resolution of 2.7Å. The measurements for the derivatives included anomalous dispersion data. The crystals were extremely stable so that a single crystal lasted for the entire week which was required for each set of X-ray measurements. Standard computational techniques were used to calculate an electron-density map. A model was then fitted to the electron-density in a Richards optical comparator, using scaled model-building components.

Note 3. (added in March, 1979). Subsequent refinement of the structure at 1.6Å resolution has produced more precise values for the positions of the protein backbone and side-chains, and has cast doubt on the description of a hydrophobic channel from Tyr 83 to the Cu atom.

References

Acland, C.B., and Freeman, H.C., 1971. *Chem. Commun.* , 1016-1017.

Acland, C.B., Flook, R.J., Freeman, H.C. and Scudder, M. L., 1972. *Acta Cryst.,* **A28**, S75-S76.

Beattie, J.K., Fensom, D.J., Freeman, H.C., Woodcock, E., Hill, H.A.O., and Stokes, A.M., 197S. *Biochim. Biophys. Acta,* **405**, 109-114.

Billo, E.J., 1974. *Inorg. Nucl. Chem. Lett.,* **10**, 613-617.

Birker, P.J.M.W.L. and Freeman, H.C., 1977. *J. Amer. Chem. Soc.,* **99**, 6890-6899.

Boulter, D., Haslett, B.G., Peacock, D., Ramshaw, J.A.M., and Scawen, M.D., 1977. *Int. Rev. Biochem.,* **13**, 1-40.

Chapman, G.V., Colman, P.M., Freeman, H.C., Guss, ,J. M. , Murata, M. , Norris, V.A. , Ramshaw, J.A.M., and Venkatappa, M.P., 1977*a*. *J. Mol. Biol.,* **110**, 187-189.

Colman, P.M., Freeman, H.C., Guss, J.M., Murata, M., Norris, V.A., Ramshaw, J.A.M., Venkatappa, M.P., and Vickery, L.E., 1977*b*. *J. Mol. Biol.*, **112**, 649-650.

Colman, P.M., Freeman, H.C., Guss, J.M., Murata, M., Norris, V.A., Ramshaw, J.A.M. and Venkatappa, M.P., 1977*c*. In *Photosynthesis 77: Proceedings of the Fourth International Congress on Photosynthesis* (Eds. Hall, D.O., Coombs, J. & Goodwin, T.W.), 809-813. (Biochemical Society, London).

Colman, P.M., Freeman, H.C., Guss, J.M., Murata, M., Norris, V.A., Ramshaw, J.A.M. and Venkatappa, M.P., 1978. *Nature,* **272**, 319-324.

Freeman, H.C., 1967. *Advanc. Prot. Chem.,* **22**, 257-424.

Freeman, H.C. Norris, V.A., Ramshaw, J.A.M., and Wright, P.E., 1978. *FEBS Lett.,* **86**, 131-135.

Gould, D.C. and Ehrenberg, A., 1968. *Eur. J. Biochem.,* **5**, 451-455.

Hare, J.W., Solomon, E.I., and Gray, H.B., 1976. *J. Amer. Chem.Soc.,* **98**, 3205-3209.

James, B.R. and Williams, R.J.P., 1961. *J. Chem. Soc.*, 2007-2019.

Jones, T.E. Rorabacher, D.B. and Ochrymowicz, L.A., 1975. *J. Amer. Chem. Soc.,* **99**, 6730-6739.

Markley, J. K. , Ulrich, F. L. , Berg, S. P. and Krogmann, D.W., 1975. *Biochemistry,* **14**, 4428-4433.

Milne, P.R. and Wells, J.R.F., 1970. *J. Biol. Chem.,* **245**, 1566-1574.

Miskowski, V. Tang, S.P.W., Spiro, T.G., Shapiro, E. and Moss, T.H., 1975. *Biochemistry,* **14**, 1244-1250.

Siiman, O., Young, N.M. and Carey, P.R., 1976. *J. Amer. Chem. Soc.,* **98**, 744-748.

Solomon, E.I., Clendening, P.J., Gray, H.B. and Grunthaner, F.J., 1975. *J. Amer. Chem. Soc.,* **97**, 3878-3879.

Solomon, E. I., Hare, J. W. and Gray, H. B., 1976. *Proc. Natl. Acad. Sci. U.S.A.,* **73**, 1389-1393.

Ugurbil, K., Norton, R.S., Allerhand, A. and Bersohn, R., 1977. *Biochemistry,* **16**, 886-894.

Vallee, B.L. and Williams, R.J.P., 1968. *Proc. Natl. Acad. Sci., USA.,* **59**, 498-505.