

Three-Dimensional Solution Structure of Plastocyanin from the Green Alga *Scenedesmus obliquus*

JONATHAN M. MOORE, DAVID A. CASE, WALTER J. CHAZIN, GARRY P. GIPPERT, TIMOTHY F. HAVEL, ROY POWLS, PETER E. WRIGHT*

The solution conformation of plastocyanin from the green alga *Scenedesmus obliquus* has been determined from distance and dihedral angle constraints derived by nuclear magnetic resonance (NMR) spectroscopy. Structures were generated with distance geometry and restrained molecular dynamics calculations. A novel molecular replacement method was also used with the same NMR constraints to generate solution structures of *S. obliquus* plastocyanin from the x-ray structure of the homologous poplar protein. *Scenedesmus obliquus* plastocyanin in solution adopts a β -barrel structure. The backbone conformation is well defined and is similar overall to that of poplar plastocyanin in the crystalline state. The distinctive acidic region of the higher plant plastocyanins, which functions as a binding site for electron transfer proteins and inorganic complexes, differs in both shape and charge in *S. obliquus* plastocyanin.

THE MECHANISMS OF LONG-RANGE biological electron transfer and the nature of the interactions between electron transfer proteins are currently topics of intense investigation. Plastocyanins, the "blue" copper proteins that function as electron carriers in photosynthetic organisms ranging from blue-green algae to higher plants, are of particular interest because of their unusual spectroscopic and electronic properties and because they exhibit well-defined binding sites for other electron transfer proteins and inorganic reagents (1, 2). To date, the x-ray structure of only one plastocyanin (Pc), from poplar leaves, has been reported (3). This structure reveals a highly distinctive negatively charged surface that functions as a binding site for cationic electron transfer complexes (4), for cytochrome c (5), and for the physiological electron transfer partner, cytochrome f (6). Most of the residues of the acidic region are altered in the Pc from the blue-green alga, *Anabaena variabilis*, and it has been suggested that the presence of an acidic patch represents a later stage in the evolutionary development of the plastocyanins (7). Three plastocyanins have been sequenced from the green algae *Chlorella fusca*, *Enteromorpha prolifera* (8), and *Scenedesmus obliquus* [reported in (1)] and, whereas these retain acidic amino acids in positions 42 to 44, the acidic patch may be disrupted by deletion of residues Met⁵⁷ and Ser⁵⁸ and substitution of Glu⁵⁹ of poplar Pc by Ala or His, and by replacement of the

highly conserved Tyr⁸³ by Phe in the proteins from *C. fusca* and *S. obliquus*. In order to examine the effect of these amino acid substitutions on the acidic patch, we have determined the conformation of the Pc from *S. obliquus*. This structure of a primitive algal Pc has additional significance, since Pc sequences have been widely used to establish phylogenetic relations for a wide range of algae and higher plants (2). A novel feature of the present structure is that it has been determined in the solution state by two-dimensional (2D) nuclear magnetic resonance (NMR) methods rather than by conventional x-ray crystallographic procedures.

Recent advances in 2D NMR spectroscopy have led to methods for the determination of the structure of small proteins in solution (9). The process of structure determination occurs in several steps, beginning with sequence-specific assignment of the NMR spectrum with sequential assignment

procedures (10) followed by evaluation of distance and dihedral angle constraints from nuclear Overhauser effects (NOE) and scalar coupling constants, and then structure determination and refinement. The most widely used method for a priori structure determination uses distance geometry algorithms to determine the spatial arrangement of atoms in a molecule (11, 12). The resulting structures may then be refined by energy minimization (13) or a combination of energy minimization and restrained molecular dynamics methods (14, 15). This approach has been used to determine the structure of *S. obliquus* Pc. In parallel, a computationally more efficient molecular replacement technique has also been used to generate solution structures by restrained molecular dynamics methods by using the x-ray crystal structure of the homologous poplar Pc as a starting structure.

Sequential resonance assignments for the backbone and most side chain protons of reduced, diamagnetic *S. obliquus* Pc (97 residues, $M_r \sim 10,000$) have been obtained with 2D NMR methods (16) and will be described in detail elsewhere. All of the NMR experiments were completed with only 15 mg of protein. NOE constraints were obtained by analysis of cross peak intensities in 2D nuclear Overhauser effect (NOESY) spectra recorded with mixing times of 50, 80, and 120 msec. Upper bounds were determined by empirical calibration versus known sequential distances (10) by a procedure similar to that used by Williamson *et al.* (13). A total of 490 distance constraints derived from NOEs was used. Dihedral angle restraints were obtained from the $^3J_{\text{NH}\alpha}$ coupling constants measured from the NH-C α H cross peaks (17) in the scalar correlated (COSY) spectrum of *S. obliquus* Pc in H₂O. The backbone torsion angle ϕ

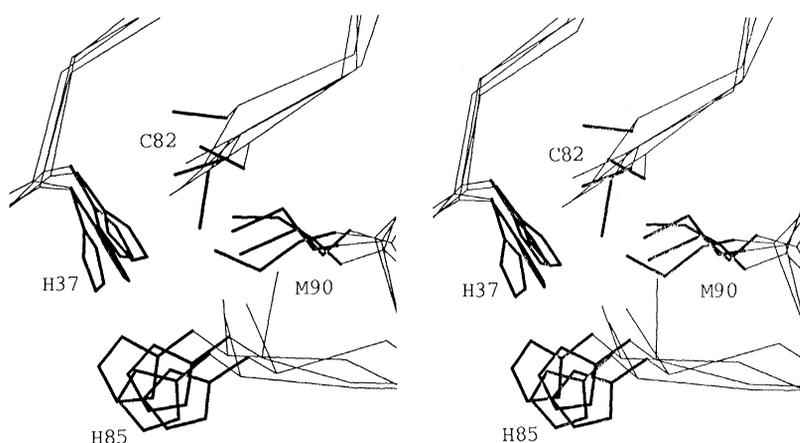


Fig. 1. Stereoview of the copper binding site in unrefined structures derived from distance geometry calculations. The copper atom has not been included in the calculations, but the relative positions of the ligand side chains are adequately defined by the observed NOE connectivities. Structures in which there are distance violations greater than 0.2 Å involving the ligand residues are omitted. Abbreviations for the amino acid residues: C, Cys; H, His; and M, Met.

J. M. Moore, D. A. Case, W. J. Chazin, G. P. Gippert, T. F. Havel, P. E. Wright, Department of Molecular Biology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, CA 92037.
R. Powls, Department of Biochemistry, University of Liverpool, Liverpool, England.

*To whom correspondence should be addressed.

was constrained (13) to $\phi = -80$ to -160 for 53 residues with $^3J_{\text{NH}\alpha} > 8$ Hz and to $\phi = -90$ to $+40$ for an additional 13 residues with $^3J_{\text{NH}\alpha} < 5.5$ Hz. For the distance geometry calculations alone, further constraints imposed by the 20 hydrogen bonds that could be identified in the β -sheets were included.

For structure determination with the molecular replacement technique, an initial structure was constructed from the x-ray structure of oxidized poplar Pc (3) by delet-

ing residues 57 and 58 and by changing the remaining amino acid side chains to those of the algal sequence. The backbone and C β carbon coordinates were held in the x-ray positions, whereas the altered side chains were constructed with idealized geometries. No attempt was made to repair the "gap" created by the deletion, or to alter side chain dihedral angles to relieve bad contacts. The dynamics refinement procedure was able to produce from this starting point a final structure as good (as measured by NMR

restraint violations and internal energy) as any obtained starting from distance geometry structures. An additional 41 starting structures were generated with the distance geometry program DISGEO (12). Pseudoatoms (18) were used wherever necessary and interproton distances were corrected accordingly. Although Pc is more than 50% larger than any protein to which the DISGEO program has heretofore been applied, the quality of the resultant structures was comparable to those reported previously (13, 19), and the computational demands are within easy reach of present computers.

The structures obtained by model-building and distance geometry approaches show significant violations of the distance constraints inferred from the NMR measurements and also have large strain energies that generally arise from nonbonded contacts shorter than the sum of the corresponding van der Waals radii. Both aspects can be substantially alleviated through "refinement" procedures that seek to minimize the sum of the computed strain energy and a penalty term derived from the distance and angle constraints implied by the NMR data. Molecular dynamics methods provide a robust means to carry out this refinement by avoiding some local minima that otherwise might trap the system.

Our procedure used the AMBER all-atom force field (20) to compute the intrinsic strain energy and added a half-parabola penalty function for proton-proton distances that violated the NMR constraints, with a force constant (14) of 40 kcal/(mol \AA^2). For nonstereospecific assignments, an r^{-6} weighting of the relevant distances was used (15). Dihedral angles falling outside the limits listed above were penalized by $K(\phi - \phi_0)^2$, where ϕ_0 is the end point of the "allowed" range and $K = 40$ kcal/(mol rad^2). The 19 best distance geometry structures, plus the model-built structure, were energy-minimized in this combined potential, then "heated" to 600 K during 2 psec using a temperature-regulated molecular dynamics algorithm (21) with a temperature relaxation time τ of 1 psec. This was followed by 2 psec of equilibration at 600 K, and an 8-psec cooling run with $\tau = 2$ psec. A final step of energy minimization produced the "refined" structures (22). Further details will be given elsewhere.

NOE measurements can provide no direct information on distances from the copper atom to protein side chains and the copper atom is not included in the distance geometry calculations. However, as a result of the many NOE constraints involving protons of the ligand side chains, the unrefined distance geometry structures clearly define the copper binding site (Fig. 1) and strongly imply

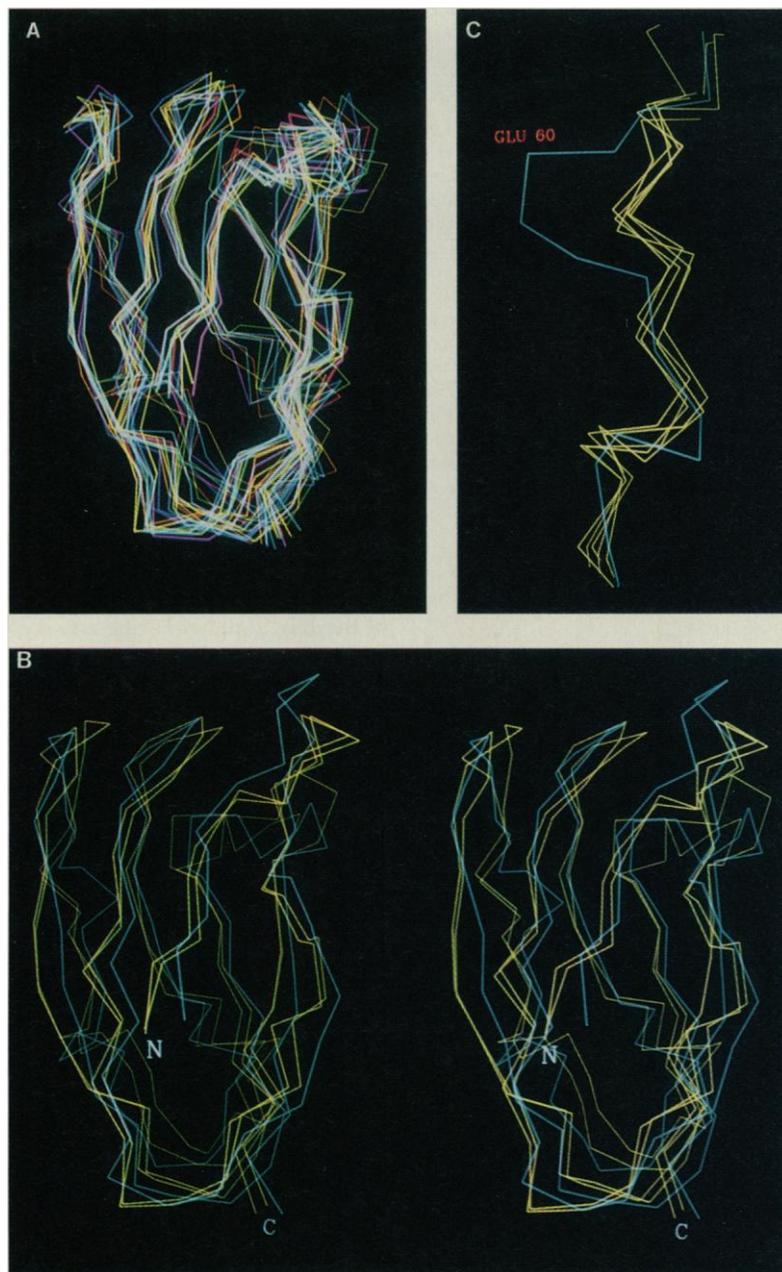


Fig. 2. (A) Superposition of nine refined structures of *S. obliquus* Pc derived from distance geometry calculations and that derived by molecular replacement. For clarity only the C α atoms are shown. The structures are superimposed for minimum rms deviation between backbone heavy atoms. (B) Stereoview of refined NMR structures of *S. obliquus* Pc (yellow) superimposed on the x-ray structure of poplar Pc (blue). The NMR structures are the best distance geometry structure (that with the lowest restraint violation energy) and that derived from the poplar Pc x-ray structure by the molecular replacement method. (C) C α chain tracing of poplar Pc x-ray structure (blue) and *S. obliquus* NMR structures (yellow) showing effects of deletions of residues 57 and 58 of the poplar Pc.

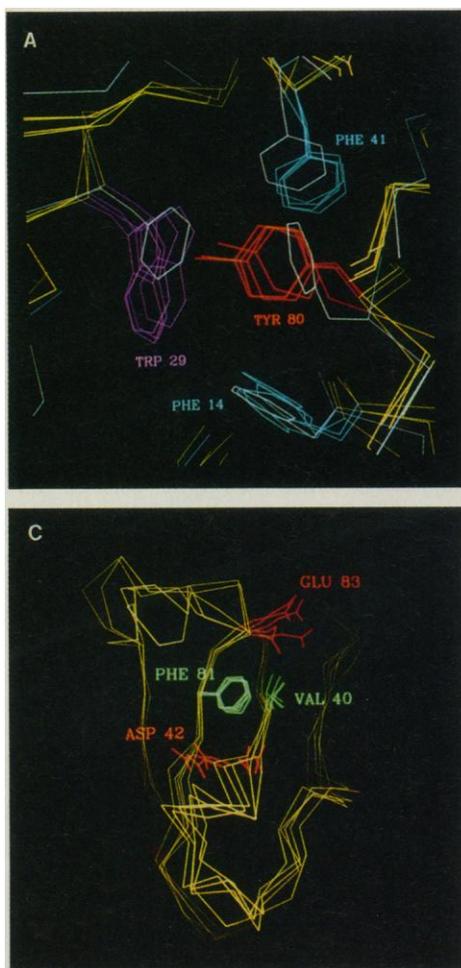


Fig. 3. (A) View of *S. obliquus* Pc showing interaction between Trp²⁹ and Tyr⁸⁰ in the interior of the β -barrel and side chains of Phe¹⁴ and Phe⁴¹. The corresponding side chains in the poplar Pc x-ray structure (Phe²⁹, Phe⁸², Phe¹⁴, and Phe⁴¹) are shown superimposed in white. The NMR and x-ray structures are superimposed for minimum rms deviation between backbone heavy atoms in the β -strands. (B) View of several surface residues in the best five NMR structures of *S. obliquus* Pc, superimposed for minimum rms deviation between all backbone heavy atoms. The side chains of His⁵⁷, Asp⁵⁸, Asp⁵⁹, Tyr⁶⁰, and Tyr⁶⁸ are shown. Tyr⁶⁸, which is not constrained by NOEs, adopts widely differing orientations. (C) View of the side chain of Phe⁸¹ and of several neighboring side chains, Val⁴⁰, Asp⁴², and Glu⁸³.

coordination by Cys⁸² S γ , Met⁹⁰ S δ , and His³⁷ N δ . Although His⁸⁵ is well situated to act as a ligand, it is not clear from the distance geometry structures alone whether coordination occurs through the N ϵ or N δ atoms. In the present case we have assumed that the copper is bound by the same donor group, that is, N δ , as in the poplar Pc x-ray structure. For the energy minimization and molecular dynamics refinement the copper atom has been inserted with bonds [force constant = 70 kcal/(mol \AA^2)] to the ligand atoms described above.

The backbone conformations of the ten best refined structures (that is, with the lowest restraint energy), which include the one obtained with the molecular replacement approach, are shown superimposed for minimum root-mean-square (rms) deviations between the backbone heavy atoms in Fig. 2A. Overall, the conformation of the polypeptide backbone is well defined by the NMR constraints with an average backbone rms deviation from the average structure of 1.2 \AA . The superimposed structures are most similar in the highly constrained β -strands (average backbone rms deviation of 0.8 \AA), but exhibit greater variability in many of the connecting loops. The greater

rms differences in the loop regions (1.5 \AA average for the backbone heavy atoms) are due, at least in part, to the smaller number of interresidue NOE constraints in these regions. The most poorly defined region of the structure is the loop between two of the copper ligands, His⁸⁵ and Met⁹⁰, where the mean backbone rms difference is 2.5 \AA . *Scenedesmus obliquus* Pc adopts a β -barrel structure made up of β -strands between residues 1 to 6, 12 to 14, 17 to 22, 25 to 31, 37 to 41, 67 to 72, 76 to 81, and 91 to 97, with a short stretch of helix between residues 52 to 56. The overall structure is similar to that of poplar Pc in the crystalline state (Fig. 2B). Some differences are apparent in the connecting loops and the helix in the *S. obliquus* Pc is shorter by one residue. The deletion of two residues (57 and 58 of the poplar Pc sequence) has the effect of eliminating a β -turn from the protein surface (Fig. 2C). This turn forms part of the negative patch in poplar Pc (3).

The conformations of most side chains within the interior of the β -barrel are well defined. These internal side chains are subject to a large number of NOE constraints as well as side chain packing constraints imposed by the global fold of the protein. Of

particular interest are Trp²⁹ and Tyr⁸⁰, both of which are replaced by Phe in the higher plant plastocyanins. The side chain conformations of these two residues are shown in Fig. 3A, superimposed upon the phenylalanines of the poplar Pc. The orientation of the Tyr⁸⁰ ring in *S. obliquus* Pc is markedly different from that of the corresponding Phe⁸² in poplar Pc. The phenol oxygen and the indole NH of Trp²⁹ are within hydrogen-bonding distance (the average O–N distance in the best ten structures is 3.0 \AA). The energetic disadvantage associated with introduction of the polar Tyr hydroxyl group into the hydrophobic interior of the protein may be compensated for by hydrogen bonding to the Trp side chain. The occurrence of the Trp²⁹-Tyr⁸⁰ pair in the *S. obliquus* and *C. fusca* plastocyanins, but not in the higher plant proteins, appears to be an example of a rare complementary mutation.

A further feature that distinguishes *S. obliquus* Pc from higher plant plastocyanins is the presence of His⁵⁷ in the algal protein. This histidine provides a defined binding site for ruthenium complexes that can be used to probe the mechanism of long-range electron transfer. The position of His⁵⁷ is well defined by the NMR constraints, despite its location on the protein surface (Fig. 3B). The structures show the imidazole ring in close contact with Asp⁵⁹, and provide an explanation for the high pK (7.8) of His⁵⁷ (23). The distance between the nearest atom of the His⁵⁷ imidazole ring and the copper atom, determined from the NMR structures, is 11 to 12 \AA , yet the rate of intramolecular electron transfer to Cu(II) from (NH₃)₅Ru(II) attached to His⁵⁷ is less than 0.26 sec⁻¹ (24). This is considerably slower than the rates of electron transfer over comparable distances in other proteins (25) or over longer distances in plastocyanins (24). It thus appears that the rate of intramolecular electron transfer is not a simple function of distance between donor and acceptor but may depend on the nature of the intervening amino acid residues.

The highly conserved Tyr⁸³ of the higher plant (and some algal) plastocyanins is replaced by Phe⁸¹ in *S. obliquus* Pc. This residue is near the binding site for cationic electron transfer reagents (4). Despite its surface location, the Phe⁸¹ side chain in *S. obliquus* Pc is constrained by NOEs to residues 40, 42, and 83, and its conformation is exceptionally well defined in the NMR structures (Fig. 3C). Interestingly, Guss and Freeman have remarked on the particularly low temperature factors for the corresponding Tyr⁸³ in the x-ray structure of poplar Pc (3). We note that the replacement of Tyr by Phe at this position had no significant effect on the electron transfer kinetics for cationic

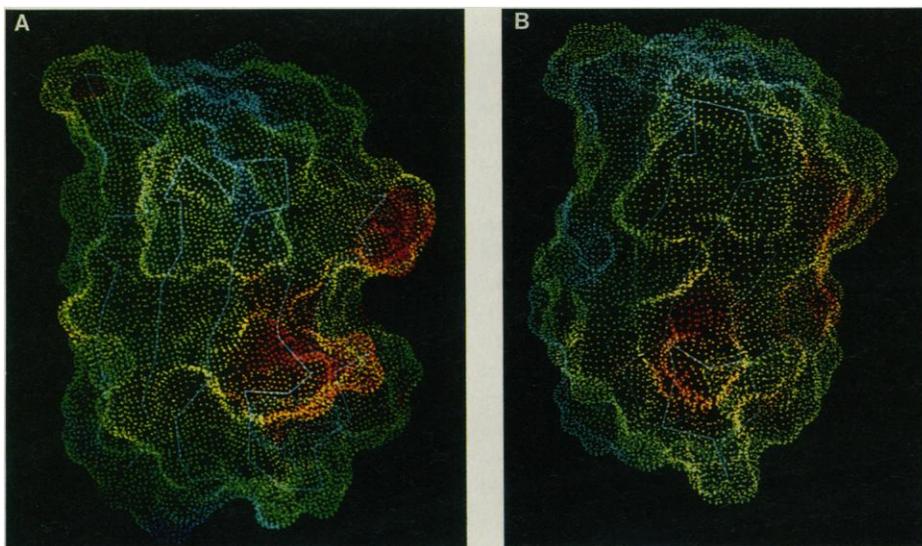


Fig. 4. Electrostatic potential at the surface of (A) the x-ray structure of poplar Pc and (B) one of the NMR structures of *S. obliquus* Pc. Note the distinctive shape of the acidic region of the poplar Pc. Surfaces were calculated with the algorithm of Connolly (27) with a 1.4 Å probe radius. The electrostatic potential was calculated by the method of Getzoff *et al.* (28) with a distance-dependent dielectric constant $\epsilon = \epsilon_0 r$, $\epsilon_0 = 4.0 \text{ \AA}^{-1}$. The surfaces are color-coded according to electrostatic potential: red, $V < -11 \text{ kcal/mol}$; yellow, $-11 \text{ kcal/mol} < V < -5 \text{ kcal/mol}$; green, $-5 \text{ kcal/mol} < V < +4 \text{ kcal/mol}$; blue, $V > +4 \text{ kcal/mol}$.

reagents (23).

The acidic side chains of residues 42 to 44, 59, and 60 of higher plant plastocyanins define a highly distinctive negatively charged surface (Fig. 4A) that functions as a binding site for cationic electron transfer reagents (4) and cytochromes c and f (5, 6). The acidic residues at positions 42 to 44 are conserved in *S. obliquus* Pc and, together with Asp⁵³, Asp⁵⁸, and Glu⁸³, form a delocalized region of negative electrostatic potential. However, the deletion of residues 57 and 58 and the substitution of Glu⁵⁹ of poplar Pc by His (His⁵⁷ in the *S. obliquus* sequence) substantially alters both the shape and charge of the acidic patch in *S. obliquus* Pc (Fig. 4B). The most notable change is the elimination as a result of the deletions of a pronounced negative protrusion from the surface of the poplar Pc. Most of the acidic residues that form the negative patch are replaced by neutral or positive side chains in the Pc from the blue-green alga, *A. variabilis*, for which there is no observable binding of cationic electron transfer complexes (26). Thus it appears that during evolution from the blue-green algae through the green algae to higher plants, the acidic patch of Pc has become progressively a distinctive structural feature that probably plays an important role in recognition and binding of the physiological electron transfer partner, cytochrome f.

REFERENCES AND NOTES

1. A. G. Sykes, *Chem. Soc. Rev.* **14**, 283 (1985).
2. D. Boulter *et al.*, *Int. Rev. Biochem.* **13**, 31 (1977).
3. J. M. Guss and H. C. Freeman, *J. Mol. Biol.* **169**,

521 (1983).

4. D. J. Cookson, M. T. Hayes, P. E. Wright, *Nature (London)* **283**, 682 (1980); *Biochim. Biophys. Acta* **591**, 162 (1980); P. M. Handford *et al.*, *J. Inorg. Biochem.* **13**, 83 (1980).
5. G. C. King *et al.*, *Biochim. Biophys. Acta* **806**, 262 (1985).
6. D. Beoku-Betts *et al.*, *Inorg. Chem.* **24**, 1677 (1984).
7. D. J. Davis *et al.*, *Plant Physiol.* **65**, 697 (1980).
8. J. Kelly and R. P. Ambler, *Biochem. J.* **143**, 681 (1974); R. J. Simpson *et al.*, *Eur. J. Biochem.* **157**,

497 (1986).

9. K. Wüthrich, *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986).
10. M. Billeter, W. Braun, K. Wüthrich, *J. Mol. Biol.* **155**, 321 (1982).
11. T. F. Havel, I. D. Kuntz, G. M. Crippen, *Bull. Math. Biol.* **45**, 665 (1983).
12. T. F. Havel and K. Wüthrich, *ibid.* **46**, 673 (1984).
13. M. P. Williamson, T. F. Havel, K. Wüthrich, *J. Mol. Biol.* **182**, 295 (1985).
14. G. M. Clore *et al.*, *EMBO J.* **5**, 2729 (1986); G. M. Clore *et al.*, *Biochemistry* **26**, 1732 (1987).
15. G. M. Clore *et al.*, *J. Mol. Biol.* **191**, 523 (1986).
16. W. J. Chazin and P. E. Wright, *Biopolymers* **26**, 973 (1987); W. J. Chazin, M. Rance, P. E. Wright, *J. Mol. Biol.*, in press.
17. D. Marion and K. Wüthrich, *Biochem. Biophys. Res. Commun.* **113**, 967 (1983).
18. T. F. Havel and K. Wüthrich, *J. Mol. Biol.* **182**, 281 (1985).
19. K. Wüthrich *et al.*, *ibid.* **169**, 949 (1983).
20. S. J. Weiner, P. A. Kollman, D. T. Nguyen, D. A. Case, *J. Comput. Chem.* **7**, 230 (1986). Computer programs were modified from AMBER 3.0 [U. C. Singh, J. Caldwell, P. A. Kolman, University of California, San Francisco (1986)].
21. H. J. C. Berendsen *et al.*, *J. Chem. Phys.* **81**, 3684 (1984).
22. Coordinates are available on request from the authors.
23. J. McGinnis *et al.*, *Inorg. Chem.*, in press.
24. A. G. Sykes, personal communication.
25. S. L. Mayo *et al.*, *Science* **233**, 948 (1986).
26. M. P. Jackman *et al.*, *J. Am. Chem. Soc.* **109**, 6443 (1987).
27. M. L. Connolly, *Science* **221**, 709 (1983).
28. E. D. Getzoff *et al.*, *Nature (London)* **306**, 287 (1983).
29. We thank L. Tennant for technical assistance, L. Harvey for preparation of the manuscript, A. Olson and A. G. Sykes for helpful discussions, and the National Institutes of Health (grant GM36643 to P.E.W. and GM38221 to T.F.H.) for financial support. Some of the calculations were performed at the San Diego Supercomputer Center.

6 January 1988; accepted 25 February 1988

Yeast HAP2 and HAP3: Transcriptional Activators in a Heteromeric Complex

STEVEN HAHN AND LEONARD GUARENTE

Transcription of the yeast *CYC1* gene (iso-1-cytochrome c) is regulated in part by the upstream activation site UAS2. Activity of UAS2 requires both the HAP2 and HAP3 activators, which bind to UAS2 in an interdependent manner. To distinguish whether these factors bound to UAS2 cooperatively or formed a complex in the absence of DNA, HAP2 and HAP3 were tagged by gene fusion to LexA and β -galactosidase, respectively, and purified through four chromatographic steps. The copurification of LexA-HAP2, HAP3 β -galactosidase, and UAS2 binding activity shows that HAP2 and HAP3 associate in the absence of DNA to form a multisubunit activation complex.

IN NUMEROUS BIOLOGICAL SYSTEMS, enzymes with catalytic functions have been found to exist as complexes bearing nonidentical subunits. For example, enzymes such as aspartate transcarbamoylase contain nonidentical subunits that separate their catalytic and regulatory functions (1). In macromolecular synthesis, enzymes that carry out catalytic functions have been found to exist as large complexes (DNA and

RNA polymerases, spliceosomes, and ribosomes). In contrast, transcriptional activators or repressors are usually not found as multisubunit complexes, but rather, as multimers (dimers or tetramers) with identical subunits (2). In prokaryotes, this oligomeric structure, which occurs, for example, in the

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.