Conformational Mimicry. 1. 1,5-Disubstituted Tetrazole Ring as a Surrogate for the Cis Amide Bond

Janusz Zabrocki,[†] G. David Smith,[‡] James B. Dunbar, Jr.,[⊥] Hiroshi Iijima,[⊥,§] and Garland R. Marshall*[,⊥]

Contribution from the Institute of Organic Chemistry, Politechnika, Lodz, Poland, Medical Foundation of Buffalo Research Institute, Buffalo, New York 14203, and Department of Pharmacology, Washington University School of Medicine, St. Louis, Missouri 63110.
Received December 21, 1987

Abstract: Assessment of the conformational implications of chemical modification is an important aspect of analogue design. A new procedure, the assessment of conformational mimicry, which determines the percentage of sterically accessible conformations for the parent compound also available to the analogue, is used to show that 88% of the conformers allowed for the cis amide bond are also available to peptides in which the amide bond is replaced by a 1,5-disubstituted tetrazole ring that locks the amide bond in the cis conformer. This analysis was made possible by the crystal structure of a cyclic dipeptide, *cyclo*[L-Phe-$\psi(CN_4)$-L-Ala], determined in this paper. The crystals of the diketopiperazine analogue are monoclinic, space group $P2_1/c$, with cell parameters a = 11.677 (1), b = 7.742 (1), c = 13.086 (1) Å; β = 93.39 (1)°; Z = 4; and D_{calcd} = 1.368 g cm^{-3}. The tetrazole ring system is planar with all five torsional angles equal to 0°. The diketopiperazine ring system is nearly planar, and the phenylalanine ring adopts the flagpole orientation over the cyclic dipeptide. A procedure for the preparation of this class of peptide analogues by synthetic routes avoiding racemization of the amino acids of the starting dipeptide is demonstrated. The tetrazole ring provides, therefore, a synthetic probe for the role of cis-trans isomerism of *N*-alkylamide bonds, such as that of proline, in molecular recognition.

Replacement of the amide bond by surrogates to enhance metabolic stability and/or probe receptor specificity has become an increasingly important topic of research[1] as the central biological role of peptides as chemical effectors becomes more understood. Proline occupies a special role among those amino acids incorporated into peptides by normal biochemical pathways as it is the only residue leading to an *N*-alkylamide bond when incorporated into a peptide. Cis-trans isomerism of the proline amide bond involving the amino group can readily be observed[2] in the NMR of proline-containing peptides. In the case of angiotensin and thyroliberin (TRH) analogues, the quantity of cis isomer in aqueous solution was correlated[3] with the biological activity. This suggested that the cis isomer might be the one bound to the receptor and responsible for the observed biological activity. Marshall et al.[4] proposed the tetrazole ring system as a peptide bond surrogate for the cis amide bond in order to lock the dipeptide analogue into a geometry corresponding to the cis isomer and outlined a synthetic approach. Recently, Yu and Johnson[5] have provided details regarding their synthetic approach to the preparation of such dipeptide analogues. A hypothetical mechanism was proposed involving the formation of a ketene amine intermediate to explain the observed racemization of the α-carbon of the N-terminal amino acid residue during formation of the tetrazole dipeptide analogue. In this paper, modification of the reaction conditions during the formation of the imidoyl chloride intermediate by the addition of quinoline[6] is demonstrated to prevent racemization. As was noted by Yu and Johnson[5] and shown herein, the replacement of the amide by the tetrazole ring makes racemization of the adjacent α-carbons more likely, especially when treated with base. Utilizing the methods outlined and avoiding basic conditions, Zabrocki and Marshall[7] have incorporated dipeptide analogues with the desired stereochemistry into biologically active peptides such as TRH, enkephalin, and bradykinin. A major concern is the degree of geometrical and steric similarity between the tetrazole ring surrogate and the cis amide bond, which will determine the ability of the surrogate to mimic the conformations available to the cis amide bond. This paper compares the geometry of the tetrazole ring analogue of a cyclic dipeptide, Phe-Ala, as determined by X-ray crystallography with the crystal structures of diketopiperazine rings in which the amide bonds of the cyclic dipeptides are forced to assume the cis conformation because of the cyclic constraint. In addition, a novel procedure for assessing conformational mimicry is used to show that approximately 88% of the conformations accessible to the cis isomer of a dipeptide are also accessible to the tetrazole analogue.

Experimental Procedures

Melting points were determined on a Thomas-Hoover capillary melting point apparatus and are uncorrected. Optical rotations were measured in a 1-dm cell on a Perkin-Elmer polarimeter (Model 271) at 589 nm (Na D line). Elemental analyses were performed by Huffman Laboratories, Inc., Wheatridge, CO, and are in good agreement with theoretical values. ^{1}H NMR spectra were recorded on a Varian XL-300 spectrometer at 300 MHz. Splitting multiplicities are given as (s) singlet, (d) doublet, (q) quartet, (t) triplet, and (m) multiplet. The chemical shifts are reported in parts per million (ppm) relative to tetramethylsilane (TMS) in $CDCl_3$ or DMSO-d_6. ^{13}C NMR was performed on the same spectrometer at 75 MHz. When either $CDCl_3$ or DMSO-d_6 was used as solvent, it also served as an internal standard at 77.0 or 39.5 ppm, respectively. FAB mass spectra were recorded on a Finnigan 3300 spectrometer equipped with a capillaritron gas gun from PHRASOR Scientific (Duarte, CA). For TLC, 250 nm silica gel GF precoated uniplates (Analtech) were used with the following solvent systems: (A) ethyl acetate/hexane, 1:2; (B) ethyl acetate/hexane, 1:3; (C) dichloromethane/acetone, 3:1; (D) chloroform/acetone, 7:1. TLC plates were developed with either iodine or chlorine followed by starch/KI spray. For flash chromatography, silica gel 60 (Merck) was used with the solvent systems given in the text. HPLC was performed on a Beckman 110A instrument using an 8SI 10-μm Radial-PAK cartridge and ethyl ace-

* To whom correspondence should be addressed.
† Institute of Organic Chemistry, Politechnika.
‡ Medical Foundation of Buffalo Research Institute.
§ Visiting scientist. Pharmaceutical Lab, Kirin Brewery, Ltd., Tokyo, Japan.
⊥ Washington University School of Medicine.

(1) For a review of amide bond surrogates, see: Spatola, A. F. In *Chemistry and Biochemistry of Amino Acids, Peptides and Proteins*; Weinstein, B., Ed.; Dekker: New York, 1983; Vol 7, p 267.
(2) Thomas, W. A.; Williams, M. K. *J. Chem. Soc., Chem. Commun.* **1972**, 994.
(3) Liakopoulou-Kyriakides, M.; Galardy, R. E. *Biochemistry* **1979**, *18*, 1952-1957.
(4) Marshall, G. R.; Humblet, C.; Van Opdenbosch, N.; Zabrocki, J. In *Peptides: Synthesis-Structure-Function; Proceedings of the Seventh American Peptide Symposium*; Rich, D. H., Gross, E., Eds.; Pierce Chemical: Rockford, IL, 1981; pp 669-672.
(5) Yu, K.-L.; Johnson, R. L. *J. Org. Chem.* **1987**, *52*, 2051-2059.
(6) Hirai, K.; Iwano, Y.; Saito, t.; Hiraoka, T.; Kishida, Y. *Tetrahedron Lett.* **1976**, *No. 16*, 1303-1306.
(7) Zabrocki, J.; Marshall, G. R., work in progress.

0002-7863/88/1510-5875$01.50/0

tate/hexane, 1:1, at a flow rate of 1 mL/min.

Z-Phe-Ala-OMe (1). The dipeptide was prepared by the mixed-anhydride procedure with isobutyl chloroformate:[8] mp 130–131 °C (lit.[9] mp 130–131 °C); $[\alpha]^{25}_D$ −24.9° (*c* 1, MeOH); TLC R_f(A) 0.31; FABMS *m/e* 385 (MH^+), calcd for $C_{21}H_{24}O_5N_2$ 384; ^{13}C NMR ($CDCl_3$) 18.98 (Ala β-C), 39.08 (Phe β-C), 48.48 (Ala α-C), 52.93 (OCH_3), 53.10 (Phe α-C), 67.57 (CH_2Ph), 127.93, 128.39, 128.56, 128.97, 129.50, 129.59, 136.49, 136.58 (Z and PhePh), 156.26 (Z C=O), 170.73 (Phe C=O), 170.16 (Ala C=O).

Z-*ambo*-Phe-ψ(CN_4)-Ala-OMe (2). To a stirred suspension of dipeptide **1** (384 mg, 1 mmol) in dry benzene (5 mL) was added crystalline PCl_5 (210 mg, 1 mmol). A clear solution was formed after a few minutes, the stirring was continued for 45 min, and a benzene solution of hydrazoic acid was added (3 mL).[10] The reaction mixture was stirred at room temperature for 2 h before being diluted with benzene (30 mL) and washed with 1 N $NaHCO_3$ (3 × 15 mL), H_2O (2 × 15 mL), and saturated NaCl solution (15 mL). When the dried (Na_2SO_4) benzene solution was evaporated and the residue purified by flash chromatography (solvent system dichloromethane/acetone, 15:1, v/v) the tetrazole derivative (250 mg, 62.3%) was isolated as an oily product containing two diastereoisomers.

2a was crystallized from the mixture with ethyl acetate/petroleum ether: mp 143–144 °C; $[\alpha]^{25}_D$ −71.7° (*c* 1, MeOH); $[\alpha]^{25}_D$ −64.0° (*c* 1, acetone); TLC R_f(B) 0.26; HPLC t_R 6.25 min; FABMS *m/e* 410 (MH^+), calcd for $C_{21}H_{23}O_4N_5$ 409; ^1H NMR ($CDCl_3$) 1.52 (d, *J* = 7.1 Hz, 3 H, Ala β-CH_3), 3.27 (dd, *J* = 9.3, 13.2 Hz, 1 H, Phe β-CH), 3.39 (dd, *J* = 6.5, 13.2 Hz, 1 H, Phe β-CH), 3.58 (s, 3 H, OCH_3), 4.79 (q, *J* = 7.1 Hz, 1 H, Ala α-CH), 4.99 (d, *J* = 12.2 Hz, 1 H, CH_2Ph), 5.07 (d, *J* = 12.2 Hz, 1 H, CH_2Ph), 5.18–5.27 (m, 1 H, Phe α-CH), 6.64 (d, *J* = 9.1 Hz, 1 H, NH), 7.0–7.35 (m, 10 H, 2 Ph); ^{13}C NMR ($CDCl_3$) 16.45 (Ala β-C), 41.25 (Phe β-C), 47.41 (Phe α-C), 53.25 (OCH_3), 55.49 (Ala α-C), 67.23 (CH_2Ph), 127.35, 127.82, 128.14, 128.41, 128.80, 129.14, 135.22, 135.71 (Z ands PhePh), 155.20, 155.39 (CN_4, Z C=O), 168.00 (Ala C=O).

2b was isolated by flash chromatography (solvent system ethyl acetate/hexane, 1:3, v/v) as an oil: $[\alpha]^{25}_D$ −33.6° (*c* 1.45, MeOH); TLC R_f(B) 0.32; HPLC t_R 5.73 min; FABMS *m/e* 410 (MH^+), calcd for $C_{21}H_{23}O_4N_5$ 409; ^1H NMR ($CDCl_3$) 1.87 (d, *J* = 7.3 Hz, 3 H, Ala β-CH_3), 3.34 (dd, *J* = 7.8, 13.9 Hz, 1 H, Phe β-CH), 3.45 (dd, *J* = 7.2, 13.9 Hz, 1 H, Phe β-CH), 3.58 (s, 3 H, OCH_3), 4.96 (d, *J* = 12.2 Hz, 1 H, CH_2Ph), 5.03 (d, *J* = 12.2 Hz, 1 H, CH_2Ph), 5.14 (q, *J* = 7.8 Hz, 1 H, Ala α-CH), 5.52–5.60 (m, 1 H, Phe α-CH), 5.64 (d, *J* = 8.3 Hz, 1 H, NH), 7.15–7.32 (m, 10 H, 2 Ph); ^{13}C NMR ($CDCl_3$) 16.72 (Ala β-C), 39.57 (Phe β-C), 46.68 (Phe α-C), 53.40 (OCH_3), 55.69 (Ala α-C), 67.48 (CH_2Ph), 127.25, 127.93, 128.33, 128.54, 128.74, 129.28, 135.39, 135.60 (Z and PhePh), 155.65, 155.84 (CN_4, Z C=O), 168.28 (Ala C=O).

Z-L-Phe-ψ(CN_4)-L-Ala-OMe (2a). To a stirred suspension of PCl_5 (1.26 g, 6 mmol) in chloroform (20 mL) was added quinoline (1.57 g, 12 mmol) at 10 °C. The mixture was stirred for 20 min before being cooled to 0 °C. The crystalline dipeptide **1** (2.30 g, 6 mmol) was added in portions with stirring. After 30 min at 0 °C, the mixture was allowed to warm to room temperature (a clear solution had formed), and a benzene solution of hydrazoic acid (18 mL) was added. The reaction mixture was stirred at room temperature for 2 h before being evaporated. The crude residue was partitioned between ethyl acetate and water (50 mL of each). The organic layer was washed with 1 N HCl (2 × 30 mL), 1 N $NaHCO_3$ (2 × 30 mL), H_2O (2 × 30 mL), and saturated NaCl solution (30 mL). When the dried (Na_2SO_4) ethyl acetate solution was evaporated and the residue purified by flash chromatography (solvent system dichloromethane/acetone, 15:1, v/v), the tetrazole derivative **2a** (660 mg, 26.9%) was isolated: mp 145–146.5 °C; $[\alpha]^{25}_D$ −65.5° (*c* 1, acetone); TLC R_f(B) 0.26; HPLC t_R 6.25 min; FABMS *m/e* 410 (MH^+) calcd for $C_{21}H_{23}O_4N_5$ 409.

Z-D,L-Phe-ψ(CN_4)-D,L-Ala-NH_2 (3). The tetrazole dipeptide **2a** (409 mg, 1 mmol) was dissolved in methanol (10 mL), and a vigorous stream of ammonia gas was introduced while the reaction mixture was cooled in an ice–water bath. After 90 min, the ammonia was discontinued; the flask was closed, allowed to warm to room temperature, and left to stand overnight. The solvent was removed in vacuo and the crude amide **3** (346 mg, 87.8%; two spots on TLC R_f(C) 0.62 and 0.69) was used for further synthesis. An analytical sample (R_f 0.69) was crystallized from ethyl acetate: mp 219–220 °C; $[\alpha]^{25}_D$ −0.6° (*c* 1, DMF); FABMS *m/e* 395 (MH^+), calcd for $C_{20}H_{22}O_3N_6$ 394; ^1H NMR (DMSO-d_6) 1.44 (d, *J* = 6.95 Hz, 3 H, Ala β-CH_3), 3.19 (dd, *J* = 6.9, 13.6 Hz, 1 H, Phe β-CH), 3.28 (dd, *J* = 8.7, 13.5 Hz, 1 H, Phe β-CH), 4.83–5.01 (m, 3 H, Ala α-CH, CH_2Ph), 5.23–5.36 (m, 2 H, Phe α-CH, NH), 7.19–7.34 (m, 10 H, 2 Ph), 8.29 (s, 2 H, NH_2); ^{13}C NMR (DMSO-d_6) 17.32 (Ala β-C), 38.95 (Phe β-C), 46.55 (Phe α-C), 55.78 (Ala α-C), 65.59 (CH_2Ph), 126.46, 127.29, 127.54, 128.02, 128.08, 129.25, 136.46, 136.49 (Z and PhePh), 155.60, 155.65 (CN_4, Z C=O), 168.73 (Ala C=O).

***cyclo*[D,L-Phe-ψ(CN_4)-D,L-Ala] (4).** Compound **3** (300 mg, 0.76 mmol) in AcOH (20 mL) and water (2 mL) was hydrogenated in the presence of 10% Pd on charcoal (100 mg) for 6 h. The catalyst was filtered off and washed with AcOH (10 mL). The solvent was removed in vacuo to yield **4** (169 mg, 92%) as a mixture of two racemates (L-L,D-D; L-D,D-L), which were separated by flash chromatography (solvent system $CHCl_3$/acetone, 7:1, v/v).

4a (L-D,D-L): 43 mg; mp 204–205 °C; $[\alpha]^{25}_D$ −0.55° (*c* 1, DMF); TLC R_f(C) 0.66, R_f(D) 0.24; FABMS *m/e* 244 (MH^+), calcd for $C_{12}H_{13}ON_5$ 243; ^1H NMR (DMSO-d_6) 1.58 (d, *J* = 6.9 Hz, 3 H, Ala β-CH_3), 3.14–3.26 (m, 2 H, Phe β-CH_2), 4.07 (q, *J* = 6.9 Hz, 1 H, Ala α-CH), 5.42–5.43 (m, 1 H, Phe α-CH), 6.83–7.21 (m, 5 H, Phe Ph), 8.96 (s, 1 H, NH); ^{13}C NMR (DMSO-d_6) 17.13 (Ala β-C), 40.70 (Phe β-C), 48.25 (Phe α-C), 53.78 (Ala α-C) 126.98, 128.08, 129.66, 134.18 (Phe Ph), 149.38 (CN_4), 165.57 (Ala C=O).

4b (L-L,D-D): 85 mg; mp 186–187 °C $[\alpha]^{25}_D$ −0.6° (*c* 1, DMF); TLC R_f(C) 0.62, R_f(D) 0.14; FABMS *m/e* 244 (MH^+), calcd for $C_{12}H_{13}ON_5$ 243; ^1H NMR (DMSO-d_6) 0.70 (d, *J* = 7.0 Hz, 3 H, Ala β-CH_3), 3.15–3.27 (m, 2 H, Phe β-CH_2), 4.94 (q, *J* = 7.0 Hz, 1 H, Ala α-CH), 5.44 (m, 1 H, Phe α-CH), 6.75–7.18 (m, 5 H, Phe Ph), 8.94 (s, 1 H, NH); ^{13}C NMR (DMSO-d_6) 17.98 (Ala β-C), 40.49 (Phe β-C), 48.40 (Phe α-C), 53.79 (Ala α-C), 126.95, 128.01, 129.98, 134.11 (Phe Ph), 148.75 (CN_4), 165.42 (Ala C=O). Anal. Calcd for $C_{12}H_{13}ON_5$: C, 59.24; H, 5.38; N, 28.79, Found: C, 58.94; H, 5.32; N, 28.55. $[\alpha]^{25}_D$ −0.3° (neat DMF, values above are uncorrected).

Crystal Structure. Single crystals of compound **4b** were grown by slow evaporation of an ethyl acetate solution. A single crystal (0.3 × 0.6 × 1.2 mm) was mounted, and 1607 independent reflections ($(\sin\theta)/\lambda$ = 0.55 $Å^{-1}$) were collected with Ni-filtered Cu Kα radiation (1.5418 Å) on a Nicolet P3 diffractometer using θ–2θ scans. A total of 50 data were found to have $F < 6\sigma(F)$ and were therefore considered unobserved. The intensities were corrected for Lorentz and polarization factors but not for extinction or absorption ($\mu_{Cu\,K\alpha}$ = 7.76 cm^{-1}). Real and imaginary dispersion corrections were applied to the atomic scattering factors.[11] The variance of each structure factor was calculated according to the method of Stout and Jensen:[12] $\sigma^2(F) = (1/4Lp)(\sigma^2(I) + (0.06I)^2)$. $w(F) = 1/\sigma^2(F)$. Unobserved data were assigned zero weight.

The structure was solved through the use of the direct-methods program MULTAN,[13] from which the positions of all non-hydrogen atoms were determined. Isotropic refinement of the structure by full-matrix least squares, minimizing $\sum w(F_o - F_c)^2$, converged at a residual $R = \sum||F_o| - |F_c||/\sum|F_o|$ of 0.196. Because of the high value of the residual, a $\delta(R)$ probability plot[14] was calculated and suggested that 164 data had been erroneously measured. Subsequent examination of the profiles of each reflection showed that a malfunction of the shutter/software had occurred for these reflections, yielding a discontinuous profile with low background. These data were assigned weights of zero and were excluded from all calculations. The refinement was then continued by treating the vibration of all non-hydrogen atoms anisotropically. Hydrogen atom positions were determined from difference maps and were included in the refinement, treating their vibration isotropically. The refinement converged at residual 0.044 for 1393 data (R = 0.061 for all data) and a weighted residual ($R_w = \sum w(F_o - F_c)^2/\sum wF_o^2$) of 0.071. S, the standard deviation of an observation of unit weight ($S = \sum(F_o - F_c)^2/(m - n)$ where m is the number of observations and n is the number of parameters), was 2.769. A final $\delta(R)$ plot was calculated and was found to be essentially linear with an intercept of zero and a slope of 2.5, suggesting that the standard deviations may be underestimated by a factor of approximately 2.5. The atomic coordinates and the equivalent isotropic thermal parameters[15] for the non-hydrogen atoms are listed in Table I. Tables of observed and calculated structure factors, anisotropic thermal parameters, hydrogen atom coordinates, and thermal parameters are available from G.D.S. upon request.

(8) Anderson, G. W.; Zimmerman, J. E.; Callahan, F. M. *J. Am. Chem. Soc.* **1967**, *89*, 5012.
(9) Goldschmidt, S.; Gupta, K. K. *Chem. Ber.* **1965**, *98*, 2831.
(10) Von Braun, J. *Ann. Chem.* **1931**, *490*, 125.
(11) Cromer, D. R.; Waber, J. T. *International Tables for X-Ray Crystallography*; Kynoch: Birmingham, England, 1974; Vol. 4, Tables 2.2B and 2.3.1.
(12) Stout, G. H.; Jensen, L. J. *X-Ray Structure Determination*; Macmillan: New York, 1968; p 457.
(13) Germain, G.; Main, P.; Woolfson, M. M. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* **1971**, *A27*, 368–376.
(14) Abrahams, S. C.; Keve, E. V. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* **1971**, *A27*, 157–165.
(15) Willis, B. T. M.; Pryor, A. W. *Thermal Vibrations in Crystallography*; University Press: Cambridge, 1975; pp 101–192.

Table I. Atomic Coordinates ($\times 10^4$) and Equivalent Isotropic Thermal Parameters ($\times 10^2$) for *cyclo*[L-Phe-$\psi(CN_4)$-L-Ala]

atom	*x/a*	*y/b*	*z/c*	B_{iso}
C1	1344 (1)	−1124 (2)	5113 (1)	361
C1A	956 (1)	−2752 (2)	4599 (1)	388
C1B	1867 (2)	−4192 (2)	4743 (2)	462
C1D2	3309 (2)	−4023 (3)	3387 (2)	537
C1D1	3840 (2)	−2930 (3)	5042 (2)	564
C1E2	4379 (2)	−3548 (4)	3067 (2)	685
C1E1	4898 (2)	−2454 (4)	4707 (3)	731
C1G	3026 (2)	−3714 (2)	4385 (1)	423
C1Z	5159 (2)	−2759 (4)	3724 (3)	717
C2	741 (1)	−950 (2)	2995 (1)	374
C2A	1385 (1)	561 (2)	3486 (1)	385
C2B	2526 (2)	829 (3)	3009 (2)	540
N1	646 (1)	−2406 (2)	3529 (1)	417
N2	1542 (1)	286 (2)	4588 (1)	361
N3	1908 (1)	1532 (2)	5248 (1)	462
N4	1921 (1)	839 (2)	6152 (1)	500
N5	1570 (1)	−828 (2)	6096 (1)	473
O2	338 (1)	−781 (2)	2108 (1)	455

Molecular Modeling. The SYBYL (Tripos Associates, St. Louis, MO) molecular modeling software was used to extract the crystal structures of diketopiperazines from the Cambridge Crystal Structure Database[16] for comparison with the tetrazole analogue. An averaged structure (DKPAVE) was derived and was used for comparison. The linear dipeptide, acetyl-Ala-Ala-methylamide, with the amide bond between the two alanine residues in the cis conformation, and the tetrazole analogue, acetyl-Ala-$\psi(CN_4)$Ala-methylamide, were modeled with the coordinates derived from DKPAVE for the cis amide bond or from the crystal structure of the cyclic tetrazole dipeptide **4b**. SEARCH, a module in SYBYL that determines the sterically allowed conformations by systematically varying the torsional degrees of freedom by a specified increment, was used to generate a Ramachandran plot for each of the pairs of backbone torsional angles (ϕ, ψ) associated with each amino acid residue as shown in Figure 4. The rigid-geometry approximation was used with the set of scaled van der Waals radii shown by Iijima et al.[17] to reproduce the experimental crystal data for proteins and peptides. When the cis amide dipeptide model was calculated, an option to record the vectors associated with any atom pair was used to record the orientations of the C^{α}–C^{β} bond of Ala-1 with the methylamide fixed as a frame of reference. Using the same orientation of the methylamide in the tetrazole allowed the program to determine which vectors, or orientations of the Ala-1 side chain relative to the methylamide, were common to both dipeptides. Alternatively, the acetyl group was used as the fixed frame of reference, and the side chain orientation of Ala-2 was used to monitor conformational mimicry. Since the quantitative results were essentially the same, only the former will be discussed. A torsional increment of 10° was used, and a side chain vector was assumed to correspond if both the α- and β-carbon were within 0.2 Å of the coordinates of another vector.

Results and Discussion

Initially, the conversion of a protected dipeptide, Z-Phe-Ala-OMe (**1**), to the tetrazole derivative Z-Phe-$\psi(CN_4)$-Ala-OMe (**2**) and its subsequent incorporation into the tripeptide analogue of (Phe^2)-TRH, pGlu-Phe-Pro-NH_2, to give pGlu-Phe-$\psi(CN_4)$-Ala-NH_2 appeared straightforward. The reaction scheme shown in Scheme I, however, leads to a mixture of racemic products and, subsequently, the racemic cyclic dipeptide **4**, a synthetic dead end. From subsequent studies, we believe the racemization to have occurred in two stages. First, racemization of the phenylalanine center occurs during the preparation of the tetrazole from the amide as noted by Yu and Johnson[5] unless conditions are very carefully controlled. Fortunately, the desired enantiomer **2a** could be isolated by crystallization. It could also be prepared directly if quinoline was added to the PCl_5 and conditions were closely controlled. Racemization of both the alanyl and phenylalanyl chiral centers occurred during the conversion of the methyl ester to the amide using methanolic ammonia due to the increased

Scheme I

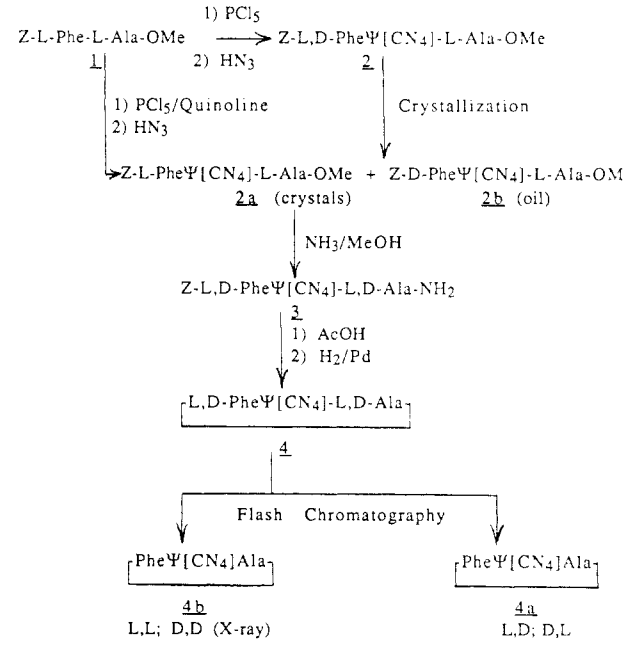

Table II. Torsion Angles (deg) for *cyclo*[L-Phe-$\psi(CN_4)$-L-Ala]

name	atoms	angle
$\psi 1$	N1–C1A–C1–N2	−10.5 (2)
$\omega 1$	C1A–C1–N2–C2A	2.9 (2)
$\phi 2$	C1–N2–C2A–C2	10.1 (2)
$\psi 2$	N2–C2A–C2—N1	−14.9 (2)
$\omega 2$	C2A–C2–N1–C1A	8.3 (3)
$\phi 1$	C2–N1–C1A–C1	4.9 (2)
	N1–C1A–C1–N5	171.2 (2)
	C1–N2–N3–N4	0.0 (2)
	N2–N3–N4–N5	0.0 (2)
	N3–N4–N5–C1	0.0 (2)
	N4–N5–C1–N2	0.0 (2)
	N5–C1–N2–N3	0.0 (2)
	C2–C2A–N2–N3	−168.4 (1)
X^1	N1–C1A–C1B–C1G	67.4 (2)
$X^{2,1}$	C1A–C1B–C1G–C1D1	−89.8 (2)
$X^{2,2}$	C1A–C1B–C1G–C1D2	89.4 (2)
	C1–N2–C2A–C2B	−113.1 (2)

acidity of the α-carbon hydrogens as a result of the adjacent tetrazole ring. The lability of the α-hydrogen of the N-terminal residue to base reflects both the tetrazole and benzyl substituents. The undesired cyclization upon removal of the α-amino protecting group was extremely facile and was essentially quantitative even when the deprotection was attempted in the presence of stoichiometric concentrations of HCl. The proximity of the α-amino group to the carbonyl of the C-terminal amide caused by the tetrazole ring made cyclization dominant. Since the synthetic justification for using dipeptides as starting materials was their availability with the desired stereochemistry and incorporation of these dipeptide analogues into peptides required free N- or C-terminals, our results suggest a need for further investigation of the deprotection conditions and/or the development of alternate synthetic routes, which are the subjects of subsequent publications.[7]

The tetrazole diketopiperazine analogue **4b** was, however, an excellent candidate for comparison with known structures of cyclic dipeptides (in which the amide bonds must assume the cis conformation to allow ring closure), allowing an assessment of the similarity between the 1,5-disubstituted tetrazole ring system and the cis amide bond. The fact that the compound crystallized in a centric space group $P2_1/c$ with cell parameters a = 11.677 (1), b = 7.742 (1), c = 13.086 (1) Å; β = 93.91 (1)°; Z = 4; and D_{calcd} = 1.368 g/cm, implied the presence of a racemic mixture in the crystal, and structural determination indicated that the crystal was composed of the L,L and D,D enantiomers. This was irrefutable proof of racemization at both chiral centers during the synthetic

(16) Crystallographic Data Center, University Chemical Laboratory, Lensfield Road, Cambridge GB2 1EW, England.
(17) Iijima, H.; Dunbar, J. B.; Marshall, G. R. *Proteins: Struct. Funct. Genet.* **1987**, *2*, 330–339.

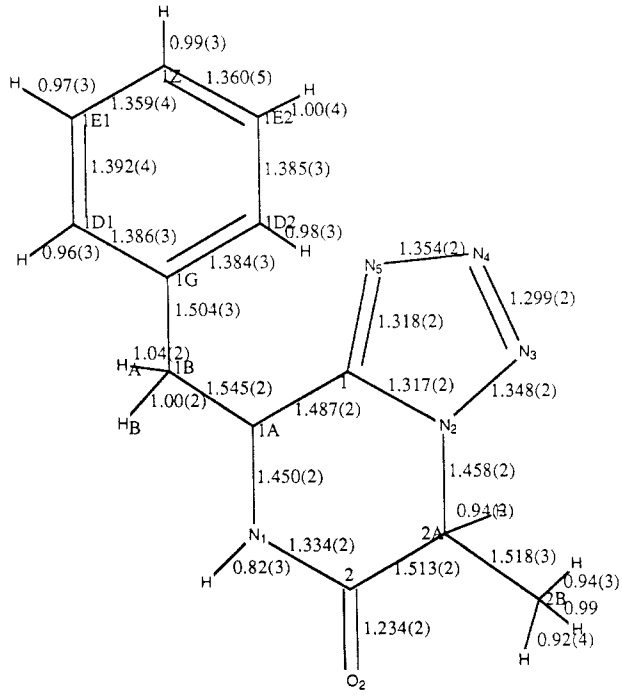

Figure 1. Structural formula, numbering scheme, and bond lengths (Å). Standard deviations are given in parentheses.

Table III. Least-Squares Equations of Selected Planes[a] and Angles (deg) between Selected Planes

plane no.	atoms	*A*	*B*	*C*	*D*	rms dev
1	N1, C1A, C1, N2, C2A, C2	0.917	−0.314	−0.245	0.156	0.065
2	C1G, C1D1, C1E1, C1Z, C1E2, C1D2	−0.368	0.895	−0.251	−5.264	0.004
3	C1, N2, N3, N4, N5	0.945	−0.293	−0.144	0.785	0.001

1–2 = 123.8; 1–3 = 6.1; 2–3 = 125.0

[a] The equation of the plane is $Ax + By + Cz = D$, where A–C are direction cosines, D is the perpendicular distance from the plane to the origin, and the rms deviation is the root-mean-square deviation of the atoms from the plane. The orthogonal coordinate system from which these values are calculated is defined by x along the a-axis, y in the ab plane, and z along the c-axis.

scheme. The data for the L,L stereoisomer with the numbering scheme, bond distances, and bond angles are illustrated in Figures 1 and 2 and are within the expected ranges.[18] Torsion angles, the equations of selected planes, and the angles between these planes are listed in Tables II and III. The overall conformation is shown in Figure 3.

The tetrazole ring system, which constrains the peptide bond to be cis, is very nearly planar. A maximum deviation of 0.030 Å from a least-squares plane is observed with the next largest deviation from the plane of 0.0001 Å. All five torsional angles are 0.0°. On the basis of bond distances, the C1–N2 bond (1.317 Å) is equivalent to the C1–N5 bond (1.318 Å), and both are shorter than the other peptide bond of the diketopiperazine (DKP) ring system (1.334 Å). The N3–N4 bond (1.299 Å) possesses the most double-bond character. The N2–N3 (1.348 Å) and N4–N5 (1.354 Å) bonds are the longest and differ by 15–20 standard deviations from the other bonds in this ring system.

The DKP ring system is best described as being nearly planar. The deviations, however, from planarity (the maximum of 0.093 Å for C2A) perturb the ring system toward a distorted-boat conformation. The angle β, which measures the degree of folding along the line joining the two α-C atoms, was calculated to be 12.1°. This value is at the upper limit of a nearly planar DKP

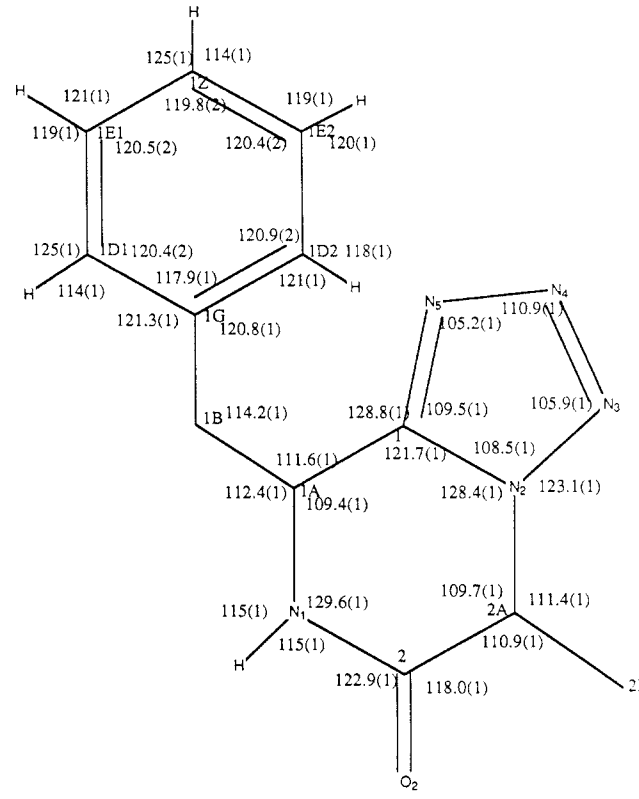

Figure 2. Bond angles (deg). Standard deviations are given in parentheses.

ring, but at the lower limit of a boat conformation.[19] The peptide bond between C1 and N2, which is part of the tetrazole ring system, deviates from planarity by 2.9°. This can be compared to the five internal torsional angles of 0.0° for the tetrazole ring system. Apparently, the preference for planarity of the tetrazole ring system is considerably greater than that of the peptide bond. It should be noted, however, that it is not uncommon for peptide bonds to deviate by 5–10° from planarity.

The phenylalanine ring is planar and adopts the familiar flagpole orientation over the DKP ring.[19] Bond distances between C1Z and the two adjacent carbon atoms, C1E1 and C1E2, are shorter than expected by approximately 0.03 Å. This shortening is most likely a result of thermal motion rather than any real structural change.

One intermolecular hydrogen bond exists between the proton of N1 and the single oxygen atom, O2, of a symmetry-related molecule [N1–O2, 2.954 (2) Å; H–O2, 2.17 (3) Å; N1–H–O2, 159 (2)°]. The net effect of this hydrogen-bonding scheme is to produce an infinite chain of diketopiperazine molecules linked together by a single intramolecular hydrogen bond and parallel to the b-axis.

As the major motivation for developing the chemistry of 1,5-disubstituted tetrazole analogues was their use as cis amide bond surrogates, an assessment of the conformational implications of their introduction into peptides was required. As one does not know in most cases the receptor-bound conformation, one can only argue statistically. Clearly, an analogue capable of assuming all the conformations available to the parent structure is more likely to be able to bind with the receptor, as it is capable of assuming the unknown receptor-bound conformation, than another analogue with only a limited conformational overlap with the parent. Some comparative measure of conformational mimicry would be useful to both interpret biological activity of analogues with amide bond modification as well as guide the chemist in deciding which analogue was worth synthetic effort. By comparison of the possible conformations available to the linear dipeptide, acetyl-L-Ala-L-

(18) Benedetti, E. In *Peptides, Proceedings of the Fifth American Peptide Symposium* Goodman, M., Meienhofer, J., Eds.; Wiley: New York, 1977; pp 257–273.

(19) Karle, I. L. In *The Peptides*; Gross, E., Meienhofer, J., Eds.; Academic: New York, 1981; Vol. 4, pp 1–54.

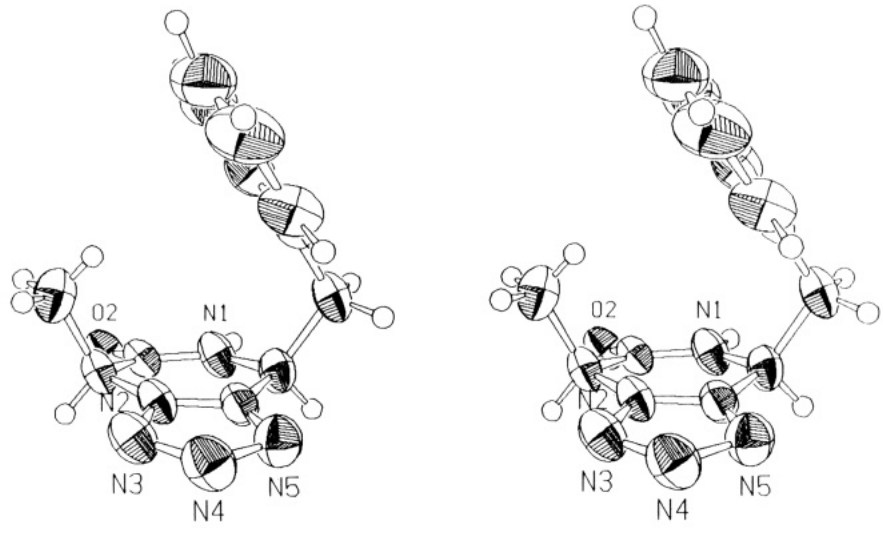

Figure 3. Stereodiagram illustrating the observed conformation. Thermal ellipsoids have been plotted at 50% probability. Hydrogen atoms have been assigned an isotropic temperature factor of 1.0 for clarity.

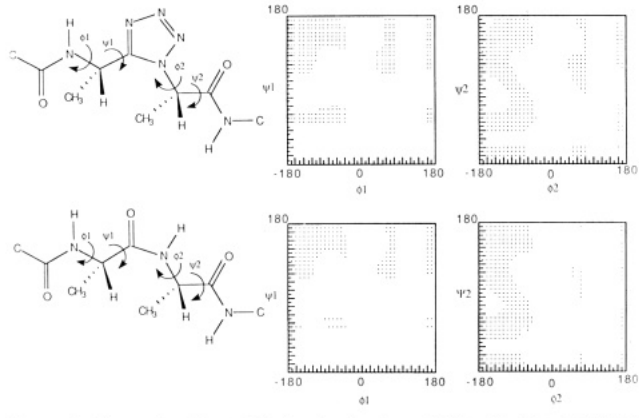

Figure 4. Ramachandran plots for torsional variables $\phi 1$, $\psi 1$ and $\phi 2$, $\psi 2$ for acetyl-L-Ala-$\psi(CN_4)$-L-Ala-methylamide (top) and acetyl-L-Ala-[cis]-L-Ala-methylamide (bottom) at 10° increments. Dots represent sterically allowed values.

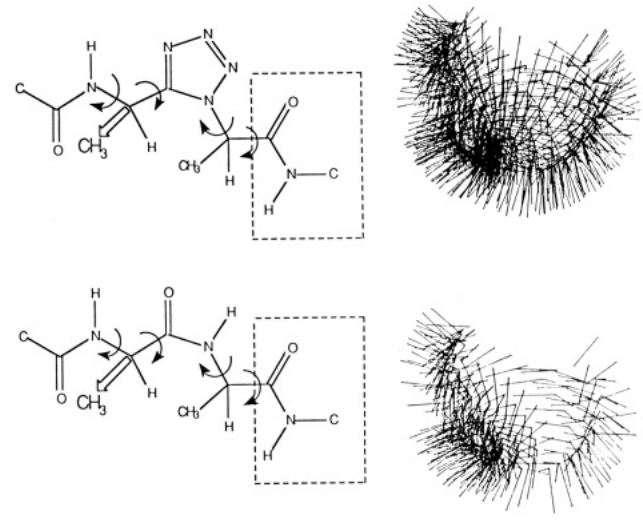

Figure 5. Vector plots of α-carbon–β-carbon bond vectors (arrows) for acetyl-L-Ala-$\psi(CN_4)$-L-Ala-methylamide (top) and acetyl-L-Ala-[cis]-L-Ala-methylamide (bottom) at 10° increments. Dashed area denotes fixed coordinates of methylamide portions used as common frame of reference.

Ala-methylamide, in which the central amide bond was fixed in the cis conformation with acetyl-L-Ala-$\psi(CN_4)$-L-Ala-methylamide in which the central amide has been replaced with the 1,5-disubstituted tetrazole ring, an assessment of the ability of tetrazole analogues to mimic the conformations accessible to cis amide conformations is made. Direct comparison of conformational maps such as the Ramachandran plots shown in Figure 4 has only limited utility as changes in molecular geometry or overall degrees of torsional freedom can make such comparisons meaningless. By directly determining all of the sterically allowed orientations of a reporter group (in this case, the side chain α–β vector) with respect to a known internal frame of reference, one can quantify those orientations available to the parent that are also available to the analogue (Figure 5). The percentage of orientations available to the analogue available to the parent is referred to as the conformational mimicry index. For the tetrazole surrogate of the cis amide bond, the conformational mimicry index is 88% (the number of vectors (747) common to both the tetrazole and cis amide divided by the total number of vectors (849) allowed for the cis amide). This is a dramatic difference from the original estimate[4] of 22% and reflects the use of the crystal data from a 1,5-disubstituted tetrazole rather than a monosubstituted tetrazole ring as well as a refined set[17] of effective van der Waals parameters that accurately reflect the sterically allowed conformations seen in protein and peptide crystal structures.

The tetrazole analogue has more conformational freedom than the cis amide model, with 33 359 conformers allowed compared to 14 912 allowed for the cis amide for 36^4 (or 1 679 616) possible conformations. This difference may be easily visualized in plots of the vector maps for the two dipeptides shown in Figure 5. The increased number of sterically allowed conformations available to the tetrazole analogues probably reflects the increase in valence angle between the C^α—C=N of the tetrazole corresponding to the C^α—C=O angle of the cis amide. This increases the number of values available to ψ^1 for the tetrazole, as can be seen in the Ramachandran plots (Figure 4). On the other hand, the increase in steric bulk in the tetrazole where the amide hydrogen is replaced by the nitrogen is clearly reflected in the Ramachandran plot as an enlarged area around $\phi^2 = -180°$, $\psi^2 = 0°$ or 180°. The relative importance of these two opposite effects determines whether the tetrazole will have more conformations available or less. In this case, the increased freedom of the torsional rotation ψ^1 dominates, and the resulting increased flexibility should be reflected in increased loss of entropy upon binding of tetrazole analogues compared with the cis amide. Since the cis amide can isomerize to the trans amide (also having multiple conformers available), entropic changes on binding are difficult to estimate. As the conformers have not been scaled energetically, only rough estimates of entropic differences can be made in any case.

One can conclude, however, that the tetrazole analogue is an excellent conformational mimic of the cis amide bond. If the role of proline, or other *N*-alkylamides found in a wide variety of biologically important peptides such as cyclosporin, is to make

the cis amide conformer energetically accessible for binding and recognition by the receptor, then the tetrazole analogue should provide a useful probe of this role. On the other hand, the cis amide offers a unique arrangement of an adjacent hydrogen-bond donor and acceptor, which the tetrazole does not possess and whose steric bulk would prevent from assuming close proximity to the receptor. Activity of analogues with tetrazole replacement will provide strong evidence for the role of the cis amide in receptor recognition. Lack of activity will not exclude the cis amide from consideration, because of the differences pointed out above. Other cis amide surrogates such as the cis olefinic isostere would be useful to probe this question. Hann et al.[20] have reported synthetic difficulties in the preparation of peptides with the cis olefinic surrogate due to rearrangement from the cis-β,γ-unsaturated isomer to the more stable trans-α,β-unsaturated isomer.

Conclusion

Conditions for the preparation of tetrazole dipeptide analogues in which the amide bond is replaced by a 1,5-disubstituted tetrazole ring have been found that preserve the chiral integrity of the starting dipeptide. A comparison of the crystal structure of a diketopiperazine in which one of the cis amides has been replaced with a tetrazole ring demonstrates clearly the strong geometric similarity of these two functional groups. A method to assess the impact of the minor geometrical changes and the increase in steric bulk on the conformational ability of the tetrazole analogue to mimic those conformers accessible to the cis amide has been developed that shows 88% mimicry. The combination of a synthetic route to chirally pure products with the knowledge that the tetrazole isostere allows nearly complete sampling of the conformers available to the cis amide bond argues strongly for its use to probe the role of proline and other *N*-alkyl amino acids in molecular recognition.

Acknowledgment. J.Z. acknowledges financial support for part of this work from the Polish Academy of Sciences (Grant CPBP 0113.2.5). The NMR spectra were obtained (J.B.D.) at the Washington University High-Resolution NMR Service Facility (NIH Grant 1S10 RR02004) and the FAB mass spectra at the Washington University Mass Spectroscopy Resource (NIH Grant RR00945). Tim Callahan assisted with retrieval from the Cambridge Database. Additional support was received from the NIH (Grant GM24483, G.R.M.; Grant GM19684, G.D.S.).

(20) Hann, M. M.; Sammes, P. G.; Kennewell, P. D.; Taylor, J. B. *J. Chem. Soc., Perkin Trans. 1* **1982**, 307-314.

Preparation and Characterization of Two His-59 Ruthenium-Modified Algal Plastocyanins and an Unusually Small Rate Constant for Ruthenium(II) → Copper(II) Intramolecular Electron Transfer over ~12 Å

M. P. Jackman,[†] J. McGinnis,[†] R. Powls,[‡] G. A. Salmon,[§] and A. G. Sykes*,[†]

Contribution from the Department of Chemistry, The University, Newcastle upon Tyne NE1 7RU, U.K., Department of Biochemistry, The University, Liverpool L69 3BX, U.K., and Cookridge Radiation Research Centre, University of Leeds, Cookridge Hospital, Leeds LS16 6QB, U.K. *Received December 10, 1987*

Abstract: Plastocyanins from the algae *Anabaena variabilis* and *Scenedesmus obliquus* possess a single uncoordinated surface histidine at position 59. Procedures for Ru modification of this residue using $[Ru(NH_3)_5H_2O]^{2+}$ are described. The modification time required is strongly dependent on the net charges on the proteins, estimated as 1+ and 9- respectively for PCu^I at pH 7. The major product in each case has been characterized by ICP atomic emission spectroscopy (1:1 ratios of Cu to Ru). The His-59 residue of the Ru-modified products no longer reacts with diethyl pyrocarbonate (DEPC). Also the sharp ^1H NMR His-59 C_2H resonance at 8.2 ppm is lost due to paramagnetic line broadening by the adjacent Ru(III). The PCu^{II}/PCu^I reduction potentials remain essentially unchanged, and the PCu^{II} UV-vis spectrum is unperturbed by Ru modification, except for the additional shoulder at 300 nm due to the $[Ru(NH_3)_5His]^{3+}$ moiety. On pulse radiolysis using $CO_2^{\bullet-}$ to reduce $PCu^{II}Ru^{III}$ (pH 7, 20 °C) the behavior observed in both cases is very similar. Reduction is partitioned between the Cu(II) (72%) and Ru(III) (28%), rate constant 6.7×10^8 M^{-1} s^{-1}, yielding stable PCu^IRu^{III} and transient $PCu^{II}Ru^{II}$, respectively. The latter decays to PCu^IRu^{III} by intra- and/or intermolecular processes (k_1, k_2), which together constitute the second stage. For *A. variabilis*, $k_1 = 0.024 \pm 0.058$ s^{-1} and $k_2 = 1.2 \times 10^5$ M^{-1} s^{-1}; and for *S. obliquus*, $k_1 = 0.04 \pm 0.22$ s^{-1} and $k_2 = 3.3 \times 10^5$ M^{-1} s^{-1}. Therefore k_1 values are <0.082 and <0.26 s^{-1}, respectively, with zero values not excluded from this study. Modification of *Pseudomonas aeruginosa* azurin was also carried out by a procedure already described. From four pulse radiolysis runs, transient $ACu^{II}Ru^{II}$ gives $k_1 = 2.5 \pm 0.8$ s^{-1} (pH 7, 17 °C), with no significant competition from k_2, in satisfactory agreement with the flash photolysis value of 1.9 ± 0.4 s^{-1} from the Gray group. Donor-acceptor distances (~12 Å) and driving forces are similar for the $PCu^{II}Ru^{II}$ and $ACu^{II}Ru^{II}$ systems. Of particular interest is the very small k_1 for both Ru-modified plastocyanins, indicating that electron transfer from the His-59 site through to the Cu is not a favorable route. On the other hand when unattached $[Ru(NH_3)_5Im]^{2+}$ is the reductant, stopped-flow studies indicate $k_{et} > 5 \times 10^3$ s^{-1} for reduction from the acidic patch (42-44) region of *S. obliquus* PCu^{II}.

The complex $[Ru(NH_3)_5H_2O]^{2+}$ is an appropriate reagent for the modification of metalloproteins because of its affinity for accessible surface histidine residues. It has for example been attached to specific histidine residues on ribonuclease A,[1] lysozyme,[2] horse heart cytochrome *c*,[3] and *Pseudomonas aeruginosa*

[†] The University, Newcastle upon Tyne.
[‡] The University, Liverpool.
[§] University of Leeds.

(1) Recchia, J.; Matthews, C. R.; Rhee, M.-J.; Horrocks, W. D., Jr. *Biochim. Biophys. Acta* **1982**, *702*, 105.
(2) Matthews, C. R.; Erikson, P. M.; Froebe, C. L. *Biochim. Biophys. Acta* **1980**, *624*, 499.

0002-7863/88/1510-5880$01.50/0