NMR studies of hydrogen bonding interactions with secondary amide and urea groups

Martin J. Deetz, Julianne E. Fahey and Bradley D. Smith*

Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556-5670, USA

ABSTRACT: The cis/trans ratios for six model secondary amides were determined by ^{1}H NMR in a range of solvent systems. The trans to cis equilibrium in chloroform is only slightly affected by addition of the hydrogen bond donor, trifluorethanol, but the cis rotamer is stabilized by an average of 0.7 kcal mol^{-1} when acetic acid is used as an intermolecular donor–acceptor template. Conversely, amide interaction with anionic hydrogen bond acceptors decreases the percentage of cis rotamer. ^{15}N NMR spectroscopy was used to determine the effect of hydrogen bonding on the trans amide structure. The direction and the magnitude of ^{15}N complex-induced-shifts indicate that both hydrogen bond donors and acceptors raise the secondary amide rotational barrier by increasing the C—N bond order. The relationship of these results to protein structure is discussed. Copyright © 2001 John Wiley & Sons, Ltd.

KEYWORDS: ^{15}N NMR; hydrogen bonding; secondary amide groups; secondary urea groups; cis/trans ratio

INTRODUCTION

Surveys of the Protein Structure Data Bank show that around 0.03% of the secondary amide bonds are assigned as cis rotamers.[1] These cis amides are usually located near the functional site in a protein[2,3] and thus play an important role in protein function. For example, the binding ability of the saccharide-binding protein concanavalin A is controlled by a calcium-induced cis/trans isomerization.[4] It is speculated that the cis bond is an energy source that can be utilized during protein function.[1,5]
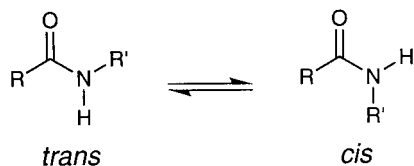
The fraction of cis amide bonds in proteins is much lower than the 1.5% cis reported for *N*-methylacetamide, a compound that has often been viewed as a model for the peptide bond.[1,5] A recent study showed that the fraction

of cis amide for dipeptides drops to around 0.4% for amino acids with C_α side chains.[6] Furthermore, peptide elongation eliminates the electrostatic stabilization and lowers the amount of cis rotamer to 0.11%. This is still almost four times higher than that seen in the data bank surveys. A possible explanation is that many cis amide bonds have been misassigned because X-ray refinement programs automatically refine amides as trans rotamers. Evidence in favor of this view is the observation that high-resolution structures (<2.0 Å) have 0.04% cis whereas low-resolution structures (>2.5 Å) have 0.01% cis.[1] Two other explanations for the scarcity of cis amides in proteins are: (a) restrictions in conformational space lower the likelihood of a cis amide; (b) non-covalent stabilization of trans amide bonds.

There is literature evidence that secondary amide conformations can be affected by non-covalent interactions. For example, protein structures show a preference for one of the amino acid residues involved in a cis amide bond to be an aromatic group, suggesting that aromatic residues are able to stabilize the cis conformation via C—H···π interactions.[5] With regard to model studies, Wolfenden and coworkers examined solvent effects and found that the fraction of cis isomer for *N*-methformamide and *N*-methylacetamide hardly changes with solvent polarity.[7] On the other hand, protonation of a formamide carbonyl is known to increase the formamide cis/trans ratio[8] and the amount of cis formamide rotamer is concentration dependent in CDCl$_3$ due to hydrogen-bonded homodimerization.[9] Cis secondary amides[10–12] and secondary carbamates[13,14] have been observed

Correspondence to: B. D. Smith, Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556-5670, USA.
Email: smith.115@nd.edu

Table 1. Percentage cis rotamer in various solvent systems at 298 K[a]

		% cis $[\Delta G(\text{kcal mol}^{-1})]$[b]					
Entry	Solvent	1	2	3	4	5	6
1	CDCl$_3$	13 [1.12]	3 [2.05]	52 [−0.05]	27 [0.59]	49 [0.01]	15 [1.02]
2	9:1 CDCl$_3$/CF$_3$CH$_2$OH	17 [0.94]	–	55 [−0.12]	27 [0.59]	–	–
3	9:1 CDCl$_3$/CH$_3$COOH	34 [0.39]	14 [1.07]	75 [−0.65]	46 [0.09]	70 [−0.50]	39 [0.26]
4	CD$_3$COOD	19 [0.86]	5 [1.74]	50 [0]	26 [0.62]	36 [0.33]	23 [0.72]
5	CDCl$_3$, 25 mM TBAA[a]	8 [1.44]	–	22 [0.75]	20 [0.82]	–	–
6	CDCl$_3$, 25 mM TBAC[a]	10 [1.30]	–	29 [0.53]	20 [0.82]	–	–

[a] [amide] = 25 mM, TBAA = tetrabutylammonium acetate, TBAC = tetrabutylammonium chloride.

[b] ΔG for trans to cis equilibrium. Error ± 5% of the stated value.

within supramolecular assemblies, but these systems have not been examined systematically.

The aim of this study is determine the effect that different hydrogen bonding motifs have on the structure of a secondary amide bond. The study is in two related parts. First, we report that the cis/trans ratios for six model secondary amides increase slightly upon addition of trifluoroethanol, a monotopic hydrogen bond donor, and increase quite significantly in the presence of acetic acid, a donor–acceptor diad. Conversely, addition of anionic hydrogen bond acceptors decreases amide cis/trans ratios. Second, we provide ^{15}N NMR evidence suggesting that hydrogen bonding to trans secondary amides increases the barrier to amide C—N rotation. The paper concludes with a short discussion on the relationship of these results to protein structure.
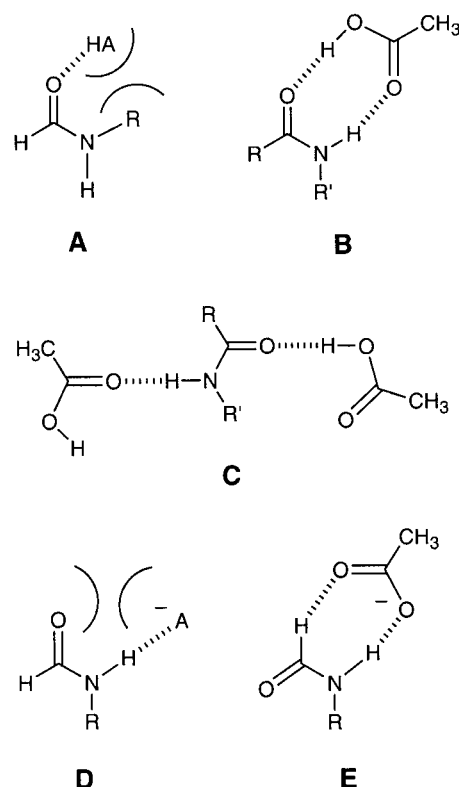
RESULTS AND DISCUSSION

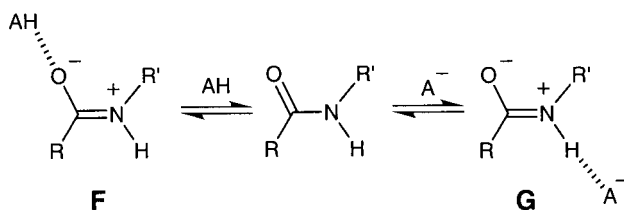### Amide cis/trans rotamer ratios

Secondary amides **1–6** were chosen for study because they have a measurable fraction of cis rotamer in CDCl$_3$
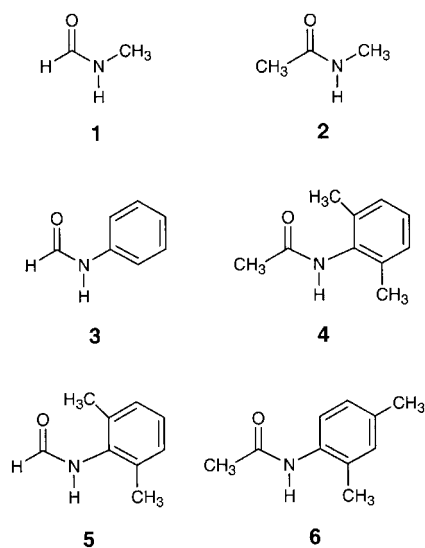
solution. Since the isomerization rate is slow on the NMR time scale, the cis/trans ratios can be determined at room temperature by direct integration. The percentages of cis rotamer found for solutions of **1–6** in various solvent systems are listed in Table 1.

The amount of cis rotamer in 9:1 CDCl$_3$/CF$_3$CH$_2$OH is slightly higher than that observed in CDCl$_3$ alone (compare entries 1 and 2 in Table 1). This is attributed to the increased steric hindrance arising in the solvated complex **A** shown in Scheme 1. A more significant increase in amide cis/trans ratio is observed when the solvent is 9:1 CDCl$_3$/CH$_3$COOH. In this case, the trans to cis equilibrium for the six amides is moved to the right by an average of 0.7 kcal mol^{-1}, presumably because the cis rotamer forms the chelated hydrogen bonded complex **B**.

Scheme 1

Scheme 2

Figure 1. Downfield change in ^{15}N chemical shift (^{15}N $\Delta\delta$) in DMSO-d_6 at 298 K for host **8** (67 mM) as a function of increasing amounts of tetrabutylammonium dihydrogen-phosphate (+) and host **9** (100 mM) with tetrabutylammonium acetate (○) and tetrabutylammonium chloride (●)

Interestingly, lower amounts of cis rotamer are observed when the solvent is 100% CD$_3$COOD (Table 1, entry 4). In this case, the concentration of acetic acid is so high that polysolvated trans amide complexes such as **C** are the predominant supramolecular structure(s).

We determined the effect of anionic hydrogen bond acceptors on secondary amide conformation by measuring the percentage of cis rotamer for amides **1, 3** and **4** in the presence of tetrabutylammonium acetate and tetrabutylammonium chloride (Table 1, entries 5 and 6). In each case, the anions decreased the percentage of cis, with acetate inducing the largest decrease. A likely explanation is that the anion forms a hydrogen bond with the amide N—H residue, which destabilizes the cis rotamer (see **D** in Scheme 1). In the case of formamides **1** and **3**, the anion may also stabilize the trans rotamer by chelating with the aldehyde C—H as shown in **E**.[15]

Hydrogen bonding and trans amide structure

^{15}N NMR spectroscopy was used to gain additional insight into the effect of hydrogen bond donors and acceptors on amide bond structure. Although ^{15}N NMR is commonly employed in biomolecule structure determination,[16,17] it is rarely employed in organic supramolecular chemistry.[18,19] The inherent insensitivity of the ^{15}N nucleus is a disadvantage, and often ^{15}N-enriched samples are required; however, this means that it is easy to monitor an isotopically enriched nitrogen compound in a mixture with other non-enriched compounds.
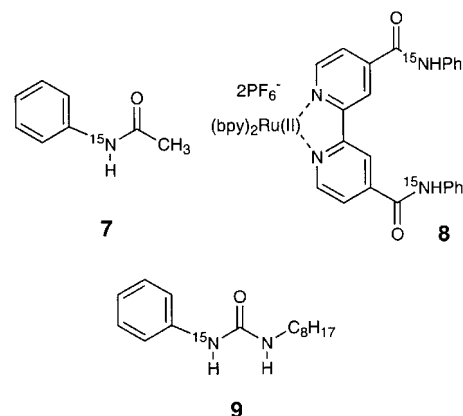
Specifically, we wished to learn how the structure of a trans secondary amide is affected by hydrogen bonding. As noted in the Introduction section, a possible reason for the scarcity of cis amides in proteins is that the trans amide is preferentially stabilized by non-covalent interactions. Initially, we examined ^{15}N-enriched acetanilide, (7) as a model amide that adopts a predominantly trans conformation. The ^{15}N chemical shift for **7** in 9:1 CDCl$_3$/CF$_3$CH$_2$OH occurs 1.8 ppm downfield of the signal in CDCl$_3$. This agrees with previous reports on the effect of hydrogen bond donors on amides and indicates that the C—N bond order increases as shown in zwitterionic complex **F** (Scheme 2).[20] The downfield movement in chemical shift is due to increased delocalization of the nitrogen lone pair, which results in increased anisotropic deshielding by the partial C=N double bond.[18,20–22]
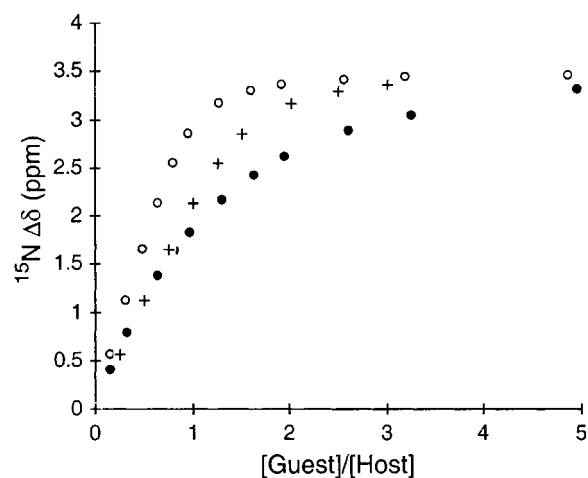
Since anionic hydrogen bond acceptors decrease cis/trans amide ratios (Table 1), it was of particular interest to determine how anions affect the structure of a trans amide. Thus, a solution of **7** in CDCl$_3$ was treated with one molar equivalent of tetrabutylammonium chloride. This resulted in a 2.8 ppm downfield movement in ^{15}N chemical shift, which again suggests an increase in C—N bond order, this time due to the zwitterionic complex **G** in Scheme 2. A factor that complicates the interpretation of this NMR result is the likelihood that **7** and chloride do not form a discrete 1:1 hydrogen-bonded complex in CDCl$_3$. Therefore, we decided to confirm this effect by using some amide-derived host compounds that are known to form 1:1 hydrogen-bonded complexes with anions.

Two ^{15}N-enriched compounds were prepared, the organometallic bis(anilide) **8**[23,24] and the urea **9**.[25] Both compounds are known to form chelated 1:1 host/guest complexes with anions, and they were used as hosts in ^{15}N NMR titration experiments. Addition of tetrabutyl-

ammonium dihydrogenphosphate to **8** in DMSO-d_6 results in a downfield movement of the amide ^{15}N resonance (Fig. 1). As expected, the titration isotherm fitted nicely to a 1:1 binding model and a value of $K_a = 110\ \text{M}^{-1}$ was extracted by iterative curve-fitting methods.[25] A downfield movement in ^{15}N chemical shift was also observed when urea derivative **9** was titrated with tetrabutylammonium acetate and tetrabutylammonium chloride in DMSO-d_6 (Fig. 1). The binding constants were calculated to be $18\ \text{M}^{-1}$ for chloride and $340\ \text{M}^{-1}$ for acetate.

As in the case of **7**, the approximate 3 ppm downfield movement of the amide (and urea) ^{15}N chemical shifts upon saturation of **8** and **9** with anions indicates that the C—N bond order increases as shown in complex **G**. Martin *et al.* have established that there is a linear correlation between C—N rotational activation energy and ^{15}N chemical shift.[26] According to their relationship, a 3 ppm downfield shift corresponds to a $0.7\ \text{kcal mol}^{-1}$ increase in activation energy for C—N rotation. In other words, forming a hydrogen bond with the NH residue of a trans secondary amide rigidifies the amide C—N bond.

CONCLUSIONS

For a protein to contain a backbone cis amide bond it must overcome an approximate $2.8\ \text{kcal mol}^{-1}$ energy difference in favor of the trans rotamer.[3,27] Our investigation of six model secondary amides in chloroform solution consistently shows that the trans to cis equilibrium is only slightly affected by the hydrogen bond donor, trifluoroethanol, but it is moved to the right by an average of $0.7\ \text{kcal mol}^{-1}$ when acetic acid (in chloroform solvent) is used as an intermolecular donor–acceptor template. Conversely, amide interaction with anionic hydrogen bond acceptors decreases the percentage of cis rotamer. It appears that if a protein is to induce one of its amide groups to adopt a cis conformation then the protein must fold in a way so as to stabilize the cis amide with at least two intramolecular hydrogen bonds in a relatively non-polar environment.

^{15}N NMR is an attractive method for studying the supramolecular chemistry of amide and urea compounds. Not only is it a useful probe for host/guest titration experiments (Fig. 1), but the direction and the magnitude of the ^{15}N complex-induced-shift provides insight into the change in C—N bond order and the corresponding C—N rotational barrier. The results of our ^{15}N NMR experiments indicate that both hydrogen bond donors and acceptors can raise amide rotational barriers by increasing the C—N bond order, as shown in structures **F** and **G** in Scheme 2.[20,28] Thus, hydrogen bonding with a protein's backbone amides is a way of rigidifying and polarizing the backbone structure.[29] The barrier for conversion of trans to cis amide is around 20 kcal mol^{-1}[5] which means that the process can be the rate-limiting

folding step for proteins containing cis amides. It is intriguing that a cis/trans isomerase for non-prolyl peptides has yet to be identified.[27]

EXPERIMENTAL

All NMR samples were 25 mM unless stated otherwise. Compounds **1–4** and **6** were purchased from Aldrich, and compound **7** was made by a literature method.[22]

2,6-Dimethylformanilide (5)

2,6-Dimethylaniline (165.1 µl, 0.67 mmol) was dissolved in benzene in a Dean–Stark apparatus. An 88% solution of formic acid (3 ml, excess) was added and the mixture was stirred under nitrogen for 30 min. The mixture was then heated to reflux for 15 h and the excess water and formic acid driven off as an azeotrope. The benzene was removed, leaving a white solid that was shown to be pure by NMR.[30] ^{1}H NMR (CDCl$_3$) δ 2.28 (s, 3H), 2.32 (s, 3H), 6.80/6.92 (bs, cis/trans NH signal, 1H), 7.12–7.15 (m, 3H), 8.11 (d, trans CHO signal, $J = 12$ Hz), 8.43 (d, cis CHO signal, $J = 1.5$ Hz) ppm.

[4,4′-Bis-(phenyl-^{15}N-carbamoyl)-2,2′-bipyridine]-bis(2,2′-bipyridine)ruthenium(II) bis(hexafluorophosphate) (8)

4,4′-Dicarboxy-2,2′-bipyridine (0.318 g, 1.3 mmol) was suspended in 15 ml of SOCl$_2$ and the mixture refluxed for 24 h under an atmosphere of nitrogen. The solvent was removed *in vacuo* and the yellow residue suspended in dry THF. ^{15}N-Aniline (263 µl, 2.86 mmol) was added via dropping funnel as a solution in THF over 5 min. The mixture was stirred at room temperature overnight under nitrogen. The mixture was filtered and the residue dissolved in dimethyl sulfoxide and precipitated upon addition of water. The precipitate was filtered and rinsed twice with water, leaving 4,4′-bis(phenyl-^{15}N-carbamoyl)-2,2′-bipyridine as an off-white solid (0.488 g, 1.23 mmol) in 95% yield. ^{1}H NMR (DMSO-d_6) δ 7.15 (t, 2H, $J = 7.5$ Hz), 7.39 (t, 4H, $J = 8$ Hz), 7.81 (d, 4H, $J = 8.5$ Hz), 7.98 (dd, 2H, $J = 5$, 1.5 Hz), 8.91 (s, 2H), 8.96 (d, 2H, $J = 5$ Hz), 10.71 (d, 2H, $J_{\text{N—H}} = 90.5$ Hz). ^{13}C NMR (DMSO-d_6) δ 163.8 ($J_{\text{N—C}} = 15.5$ Hz), 155.4, 150.0, 143.4 ($J_{\text{N—C}} = 9.5$ Hz), 138.5 ($J_{\text{N—C}} = 14.5$ Hz), 128.5, 124.1, 122.2, 120.5, 118.5. ^{15}N NMR (DMSO-d_6) (referenced relative to CH$_3$NO$_2$ at 0 ppm) -246.9. MS (FAB^{+}, NBA) $[M + \text{H}]^{+}$ at m/z 397.1449 (calc. 397.1450).

Complex **8** was formed by dissolving the above solid (0.227 g, 0.57 mmol) in a mixture of 25 ml of water, 25 ml of ethanol, and 12.5 ml of glacial acetic acid. This mixture was heated to 50 °C for 20 min under an

atmosphere of nitrogen. *cis*-Ru(bpy)$_2$Cl$_2\cdot n$H$_2$O (0.328 g, 0.63 mmol) was added and the mixture was heated at reflux for 3 days under nitrogen. The mixture was cooled to room temperature and filtered through a Celite plug, and upon removal of the solvent *in vacuo* gave a dark brown–orange glass. Column chromatography on alumina (10% methanol in CHCl$_3$ → 100% methanol) produced **8** as the dichloride salt (0.44 g, 0.54 mmol) in 95% yield. The bis-PF$_6$ salt was formed by precipitation after addition of NH$_4$PF$_6$(aq). ^{1}H NMR (DMSO-d_6) δ 7.18 (t, 2H, $J = 7.5$ Hz), 7.41 (t, 4H, $J = 8$ Hz), 7.54–7.59 (m, 4H), 7.75–7.77 (m, 6H), 7.81 (d, 2H, $J = 5.5$ Hz), 7.94 (d, 2H, $J = 6$ Hz), 7.99 (d, 2H, $J = 6$ Hz), 8.20 (d, 2H, $J = 6.5$ Hz), 8.23 (d, 2H, $J = 8$ Hz), 8.87 (d, 4H, $J = 8$ Hz), 9.39 (s, 2H), 10.77 (d, 2H, $J_{N—C} = 90$ Hz). ^{13}C NMR (CD$_3$CN) δ 163.0 ($J_{N—C} = 17$ Hz), 158.6, 157.9 ($J_{N—C} = 18$ Hz), 153.6, 152.8 ($J_{N—C} = 17$ Hz), 144.0 ($J_{N—C} = 9.5$ Hz), 139.3 ($J_{N—C} = 4.5$ Hz), 139.0 ($J_{N—C} = 14$ Hz), 130.0, 128.8 ($J_{N—C} = 6.5$ Hz), 126.5, 126.1, 125.5, 123.4, 121.8. ^{15}N NMR (CD$_3$CN) (referenced relative to CH$_3$NO$_2$ at 0 ppm) -246.6 ($J_{N—H} = 90.5$ Hz). MS (FAB$^+$, NBA) $[M - \text{PF}_6]^+$ at *m/z* 955.1447 (calc. 955.1431).

^{15}N-Phenyl-N'-octyl urea (9)

^{15}N-Aniline (100 µl, 1.14 mmol) was dissolved in CH$_3$CN (7 ml). Octyl isocyanate (242 µl, 1.37 mmol) was added and the mixture heated to reflux for 24 h under an atmosphere of nitrogen. The mixture was cooled to room temperature and placed in a freezer for 20 min. A white solid was collected by suction filtration and rinsed with cold CH$_3$CN, leaving pure **9** as a white solid (0.251 g, 1.01 mmol) in 88% yield. ^{1}H NMR (DMSO-d_6) δ 0.85 (t, 3H, $J = 7$ Hz), 1.23–1.42 (m, 12H), 3.04 (q, 2H, $J = 7$ Hz), 6.08 (bt, 1H, $J = 5.5$ Hz), 6.85 (t, 1H, $J = 7$ Hz), 7.19 (t, 2H, $J = 8$ Hz), 7.36 (d, 2H, $J = 8.5$ Hz), 8.35 (d, 1H, $J_{N—H} = 88.5$ Hz). ^{13}C NMR (DMSO-d_6) δ 155.1 ($J_{N—C} = 20$ Hz), 140.6 ($J_{N—C} = 16$ Hz), 128.6, 120.8, 117.5, 39.0, 31.2, 29.7, 28.7, 28.7, 26.4, 22.1, 13.9. ^{15}N NMR (DMSO-d_6) (referenced relative to CH$_3$NO$_2$ at 0 ppm) -270.9 ($J_{N—H} = 88.0$ Hz). MS (FAB$^+$, NBA) $[M + \text{H}]^+$ at *m/z* 250.1945 (calc. 250.1937).

Acknowledgements

REFERENCES

1. Weiss MS, Jabs A, Hilgenfeld R. *Nat. Struct. Biol.* 1998; **5**: 676.
2. Herzberg O, Moult J. *Protein: Struct. Funct. Genet.* 1991; **11**: 223–229.
3. Banerjee S, Shiematsu N, Pannell LK, Ruvinov S, Orban J, Scharz F, Herberg O. *Biochemistry* 1997; **36**: 10857–10866.
4. Bouckaert J, Poortmans F, Wyns L, Loris R. *J. Biol. Chem.* 1996; **271**: 16144–16150.
5. Jabs A, Weiss MS, Hilgenfeld R. *J. Mol. Biol.* 1999; **286**: 291–304.
6. Scherer G, Kramer ML, Schutkowski M, Reimer U, Fischer G. *J. Am. Chem. Soc.* 1998; **120**: 5568–5574.
7. Radzicka A, Pederson L, Wolfenden R. *Biochemistry* 1988; **27**: 4538–4541.
8. Stewart WE, Siddall TH. *Chem. Rev.* 1970; **70**: 517–551.
9. Manea VP, Wilson KJ, Cable JR. *J. Am. Chem. Soc.* 1997; **119**: 2033–2039.
10. Perní GJ, Kilburn JD, Essex JW, Mortshire-Smith RJ, Rowley M. *J. Am. Chem. Soc.* 1996; **118**: 10 220–10 227.
11. Bairaktari E, Mierke DF, Mammi S, Peggion E. *J. Am. Chem. Soc.* 1990; **112**: 5383.
12. Beijer FH, Sijbesma RP, Vekemans JAJM, Meijer EW, Kooijman H, Spek AL. *J. Org. Chem.* 1996; **61**: 6371–6380.
13. Marcovici-Mizrahi D, Gottlieb HE, Marks V, Nudelman A. *J. Org. Chem.* 1996; **61**: 8402–8406.
14. Moraczewski AL, Banaszynski LA, From AM, White CA, Smith BD. *J. Org. Chem.* 1998; **63**: 7258–7262.
15. Corey EJ, Barnes Seeman D, Lee TW. *Tetrahedron Lett.* 1997; **38**: 1699–1702.
16. Von Philipson W, Müller R. *Angew. Chem. Int. Ed. Engl.* 1986; **25**: 383–413.
17. Cordier F, Crzesiek S. *J. Am. Chem. Soc.* 1999; **121**: 1601–1602.
18. Haushalter KA, Lau J, Roberts JD. *J. Am. Chem. Soc.* 1996; **118**: 8891–8896.
19. Dega-Szafran Z, Szafran M, Sitkowski J, Stefaniak L. *J. Phys. Org. Chem.* 1996; **9**: 746–750.
20. Saito H, Tanake Y, Nukada K. *J. Am. Chem. Soc.* 1971; **93**: 1077–1081.
21. Linás M, Horsley WJ, Klein MP. *J. Am. Chem. Soc.* 1976; **98**: 7554–7558.
22. Lumsden MD, Wasylishen RE, Eichele K, Schindler M, Penner GH, Power WP, Cutirs RD. *J. Am. Chem. Soc.* 1994; **116**: 1403–1411.
23. Beer PD, Szemes F, Balzani V, Salà CM, Drew MGB, Dent SW, Maestri M. *J. Am. Chem. Soc.* 1997; **119**: 11864–11875.
24. Deetz MJ, Smith BD. *Tetrahedron Lett.* 1998; **39**: 6841–6844.
25. Hughes MP, Smith BD. *J. Org. Chem.* 1997; **62**: 4492–4501.
26. Martin GJ, Gouesnard JP, Dorie J, Rabiller C, Martin ML. *J. Am. Chem. Soc.* 1977; **99**: 1381–1384.
27. Scholz C, Scherer G, Mayr L, Schindler T, Fischer G, Schmid FX. *Biol. Chem.* 1998; **379**: 361–365.
28. Suezawa H, Tsuchiya K, Tahara E, Hirota M. *Bull. Chem. Soc. Jpn.* 1988; **61**: 5047–4062.
29. Scheiner S, Kern CW. *J. Am. Chem. Soc.* 1977; **99**: 7042–7050.
30. Deslongchamps P, Gerval P, Cheriyan UO, Guida A, Tailefer RJ. *Nouv. J. Chim.* 1978; **2**: 631–636.