Methyl Substituent Effects for Methyltryptophans in ^{13}C nmr [1]

Minsu Lee and Robert S. Phillips*

Departments of Chemistry and Biochemistry,
University of Georgia, Athens, GA 30602

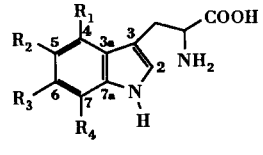
Received March 23, 1992

The ^{13}C and ^{1}H nmr spectra of methyltryptophans **2-5** in 0.1 *N* sodium deuteroxide methanol-d_4 were assigned based on 1-D and 2-D nmr techniques, including COSY, inverse-detected direct (HMQC) and long-range (HMBC) correlation. Methyl substituent effects in chemical shifts (SCS) for the indole ring of tryptophan were calculated and compared with those of indole. The correlations were linear except for 4-methyltryptophan, which suggest structural changes in the indole ring of 4-methyltryptophan and 4-methylindole. The results of molecular modeling and NOE experiments supported that suggestion.

J. Heterocyclic Chem., **29**, 1181 (1992).

Introduction.

Tryptophan is an essential amino acid containing indole, which is highly hydrophobic in nature and often occupies critical positions in proteins [2-4]. Tryptophan also exhibits unique spectroscopic properties, which enables it to serve as an intrinsic probe for protein structures, protein dynamics, and intermolecular interactions between proteins and other molecules [5-6]. Despite the extensive studies of tryptophan, its ^{13}C nmr spectra have not been fully assigned until recently [7]. Only a few studies of substituted tryptophans are known, in which the ^{13}C chemical shifts (CS) of hydroxytryptophans were reported [8-9]. In another study, we have analyzed the ^{13}C CS value of mono- and difluorotryptophans, the substituent chemical shifts (SCS) for fluorine group was calculated and their additivity was tested [1]. Despite the fact that methyltryptophans were one of the earliest derivatives of tryptophan synthesized [10-14], ^{13}C CS values of them have not been reported, even in a recent synthesis of 5- and 6-methyltryptophans [12-14]. In the present paper we unambiguously assigned ^{1}H and ^{13}C nmr spectra of **1-5**, calculated the methyl SCS, and compared the methyl SCS of tryptophan with those of methylindoles.



	R_1	R_2	R_3	R_4
1	H	H	H	H
2	CH_3	H	H	H
3	H	CH_3	H	H
4	H	H	CH_3	H
5	H	H	H	CH_3

Results.

Since the solubility of most methyltryptophans in methanol was very low (< 0.1%), the spectra were measured in 0.1N sodium deuteroxide in methanol-d_4. HMQC (heteronuclear multiple quantum coherence) [15] and HMBC (heteronuclear multiple bond correlation) [16] experiments were used instead of HETCOR [17] and COLOC [18] because of the low concentration of **1-5** (20 m*M*-25 m*M*).

Tryptophan (**1**).

The ^{1}H nmr spectra of **1** in 0.1*N* sodium deuteroxide in methanol-d_4 was almost identical to that in other solvents [19-21] and resonances were assigned readily. Only the α- and β-proton resonances showed the effect of basicity, which were shifted upfield from those in neutral methanol-d_4 [1].

There were only 7 peaks in the aromatic region of the decoupled ^{13}C spectrum of **1**. Among them, the peak at 119.7 ppm was 1.5 times higher than the other protonated carbon peaks, which indicates two carbons. These were assigned to C-4 and C-5 because their CS were 119.4 and 120.1 ppm in methanol-d_4, respectively [1]. The acquisition of direct ^{1}H-^{13}C correlation by HMQC confirmed the assignments [22]. Compared with ^{13}C spectra in methanol-d_4, C-2 moved 0.6 ppm upfield and C-3 moved 1.9 ppm downfield, which showed the enamine character of C-2 and C-3. On the other hand, C-3a moved 0.6 ppm downfield and C-7a moved 0.2 ppm upfield.

4-Methyltryptophan (**2**).

The ^{1}H nmr spectrum of **2** consists of two groups of peaks; the aromatic protons between 7-7.8 ppm and α-, β- and methyl protons between 2.7-3.6 ppm. At 300 MHz in 0.1*N* sodium deuteroxide in methanol-d_4, all the aromatic resonances were resolved. It was expected they should be one singlet and three doublets of doublets. The broad singlet at 7.1 ppm was assigned to H-2. The doublets of doublets at 6.94 ppm was attribute to H-6 since the coupling constants were 7.1 and 8.1 Hz. The assignments of H-5 and H-7 of **2** were based on corresponding CS of **1**. The higher field

signals (6.73 ppm) belong to H-5. They were not the expected doublet of doublets, but doublets of triplets, and coupling constants were 7.1, 0.9 and 0.8 Hz. On the other hand, the lower field peaks at 7.16 were doublets of doublets of doublets with coupling constants of 8.1, 0.9 and 0.4 Hz. The only logical explanation for these additional splittings was that there were long-range couplings with the methyl group. This was confirmed by selective decoupling and COSY [23] experiments shown in Figure 1.

The proton decoupled ^{13}C spectra of **2** showed 8 peaks in the aromatic region. Among protonated carbons, the lowest and the highest field signals belong to C-2 and C-7 since the corresponding CS of **1** was nearly the same. The assignments of two signals at 121.4 and 122.3 ppm were not trivial. The ambiguity was removed by HMQC spectra which showed a crosspeak between H-6 and the signal at 122.3 ppm [22]. The assignments of quaternary carbons were based on the corresponding CS of **1** and methyl SCS of indole [24-25]. Despite the 4 ppm difference, unambiguous assignment of the ^{13}C nmr signals of C-3a and C-4 was achieved by HMBC. In the 2D spectra, C-4 had a crosspeak with H-6; on the other hand, C-3a had three crosspeaks with H-2, H-5 and H-7 [22].

5-Methyltryptophan (**3**).

The aromatic ^{1}H nmr spectrum of **3** was easy to assign, due to the distinct J_{HH} splitting, and the order of CS agreed with that of 5-methylindole [26]. The highest field peak at 7.51 ppm belongs to H-4 and was

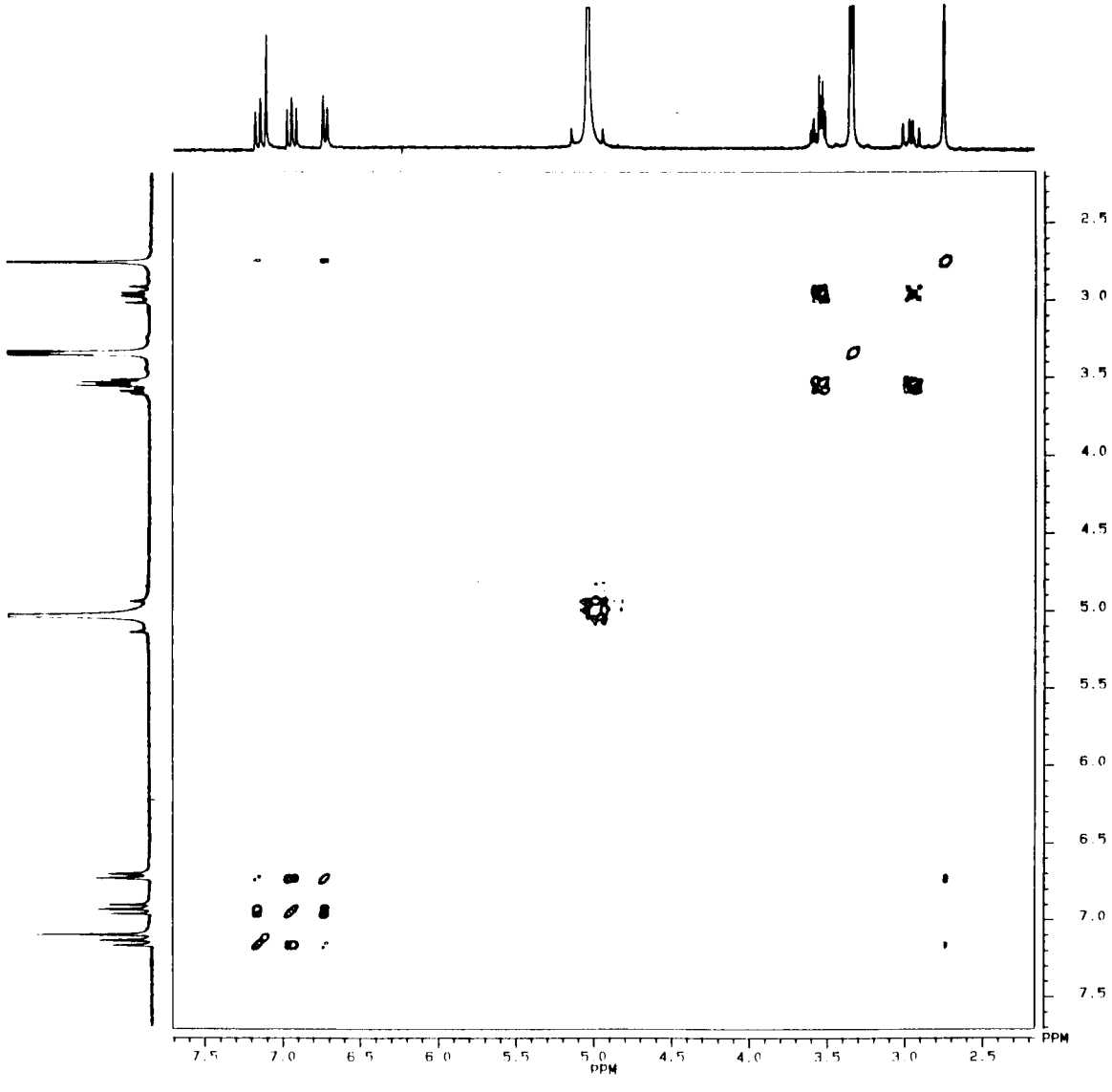


Figure 1. Contour plot of the COSY spectrum of 4-methyltryptophan.

expected to be a doublet. Instead, there were triplets with coupling constants of 1.4 and 0.8 Hz, which suggest the long range coupling with the methyl group. This was confirmed by selective decoupling and COSY experiments [22]. On the other hand, the lowest field signal at 6.94 ppm, attributed to H-6, exhibited a doublet with a coupling constant of 8.2 Hz, which indicate it had no detectable coupling with methyl group. It was interesting to see that only one of the *ortho* protons coupled with the methyl group since they are equally distant from it.

The HMQC spectrum of **3** made it possible to assign the resonances of the protonated carbons without ambiguity, including the C-6 and C-2 resonances, which appeared to be less than 1 ppm apart in ^{13}C nmr spectra [22]. Among quaternary carbon signals, two signals at 128.7 and 129.3 ppm should be either C-5 or C-3a since C-3 and C-7a resonances were easy to assign due to their characteristic CS, even though the C-3 signal happened to be together with C-7. When Parker and Roberts examined methyl SCS of methylindoles, they observed that the signals of methylated carbons were always higher than other quaternary carbons [24]. Because the signal at 128.7 ppm was higher than others, it was assigned to C-5 and the other at 129.3 ppm was assigned to C-3a (Table 2). These assignments were confirmed by HMBC experiments, as shown in Figure 2. In the 2D spectrum, C-5 had a crosspeak with H-7, while C-3a had two crosspeaks with H-2 and H-7.

6-Methyltryptophan (**4**).

The COSY spectra of **4** exhibited the expected couplings between aromatic protons [22]. Like **3**, there were long range coupling of 0.7 Hz between the methyl group and only one of *ortho* protons, which is H-7. The order of increasing CS of aromatic protons was H-4 > H-7 > H-2 > H-5 as expected.

The correlation of these assignments with ^{13}C nmr signals in the HMQC spectrum made it possible to assign the ^{13}C resonances of protonated carbons. H-4 and H-5 exhibited correlation with the carbons at 119.4 and 121.4, as illustrated in Figure 3. For the assignments of signals of quaternary carbons, we followed the same strategy used in the case of **3**, which assign C-3 and C-7a resonances, then assign the methylated carbon by comparison of height of peaks and finally confirm the assignments by HMBC [22]. In

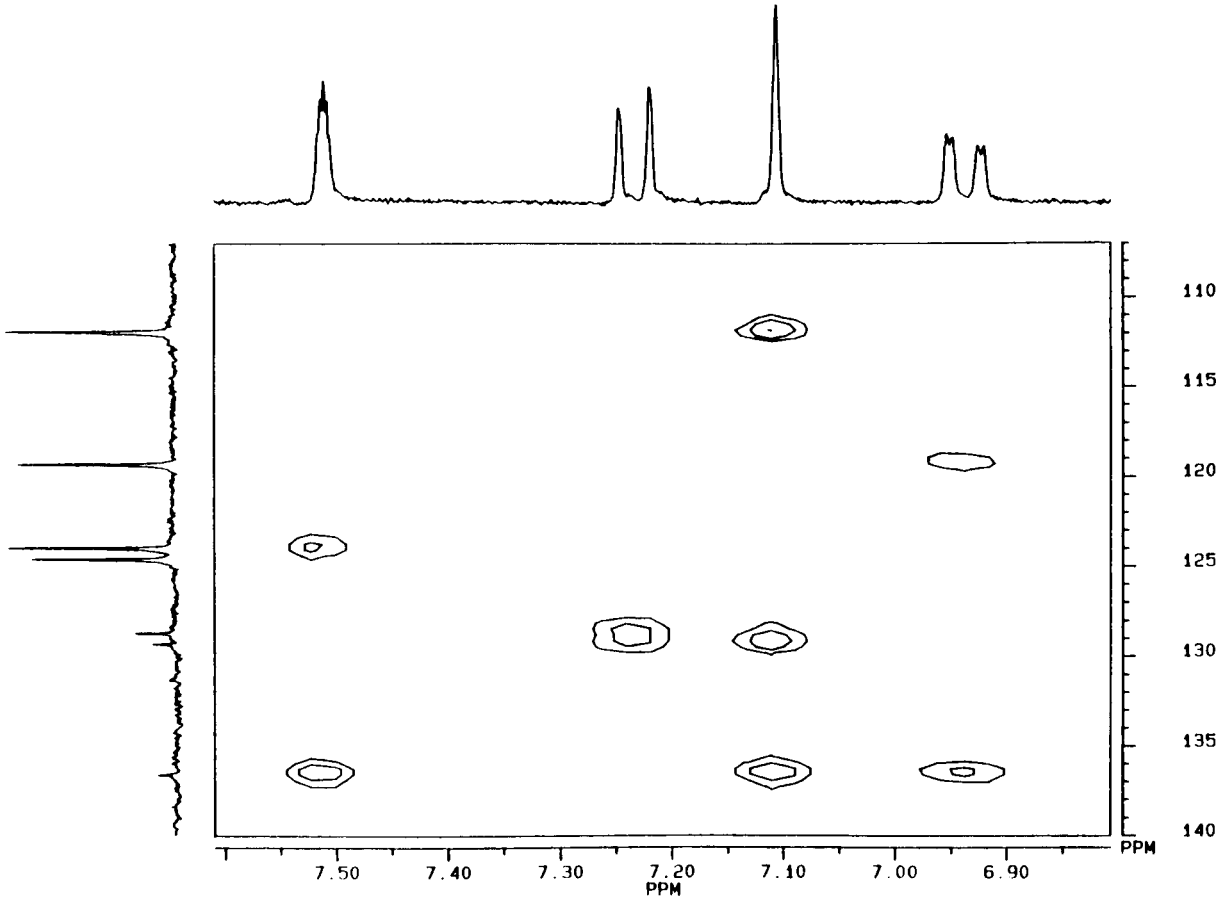


Figure 2. Portion of the 2D ^{1}H-^{13}C HMBC spectrum of 5-methyltryptophan.

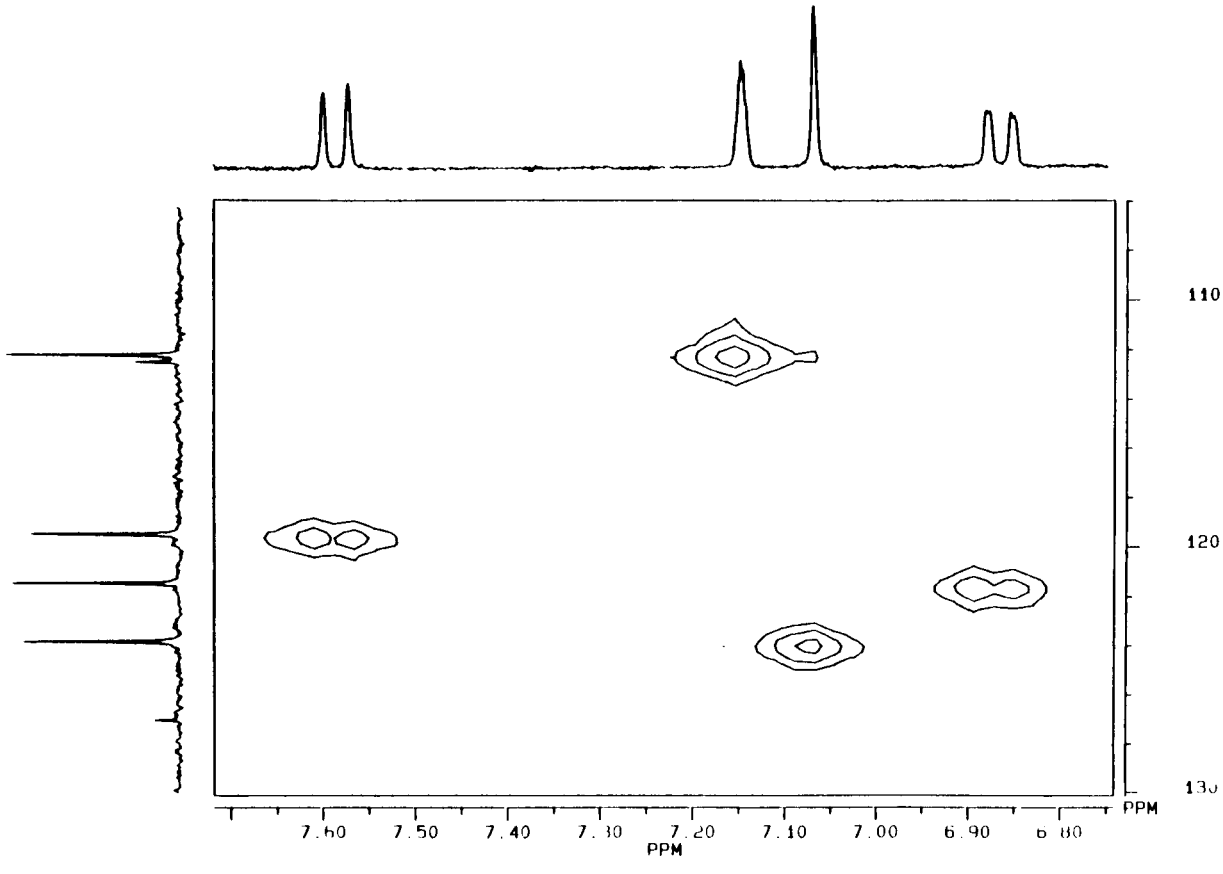


Figure 3. Portion of the 2D ^{1}H-^{13}C HMQC spectrum of 6-methyltryptophan.

Table 1

^{1}H chemical shifts of methyltryptophans

Compound	H-2	H-4	H-5	H-6	H-7	a-H	b-H	b-H'	CH_3
1	7.14	7.71	7.00	7.09	7.33	3.57	3.32	2.91	-
2	7.10	-	6.73	6.94	7.16	3.53	3.55	2.97	2.73
3	7.09	7.51	-	6.94	7.25	3.57	3.32	2.85	2.43
4	7.06	7.58	6.86	-	7.14	3.56	3.31	2.88	2.43
5	7.16	7.54	6.94	6.90	-	3.57	3.31	2.90	2.49

Table 2

^{13}C chemical shifts of methyltryptophans

Compound	C-2	C-3	C-3a	C-4	C-5	C-6	C-7	C-7a
1	124.5	112.5	129.1	119.7	119.7	122.3	112.2	138.2
2	124.5	113.6	127.2	131.3	121.4	122.3	110.2	138.7
3	124.6	111.9	129.3	119.3	128.7	123.9	111.9	136.6
4	124.8	112.3	127.0	119.4	121.4	131.9	112.0	138.7
5	124.3	112.9	128.7	117.4	119.9	122.9	121.6	137.6

the HMBC spectrum, C-3a had one long range correlation with H-4, while C-6 had three long range correlations with H-7, H-2 and H-5. The order of ^{13}C resonances agreed with that of 6-methylindole [25].

7-Methyltryptophan (**5**)

The ^{1}H nmr spectrum of **5** in the aromatic region consists of doublets of doublets of doublets at 7.54 ppm, a broad singlet at 7.16 ppm, a triplet at 6.94 ppm, and doublets of multiplets centered at 6.9 ppm. A joint utilization of the COSY spectrum [22] and CS criteria of **1** led to the assignment of all resonances; 7.54 ppm for H-4, 7.16 ppm for H-2, 6.94 ppm for H-5 and 6.9 ppm for H-6. H-4 was coupled to H-5, H-6 and the methyl group with the coupling constants of 8.0, 1.2 and 0.4 Hz, respectively. On the other hand, H-6 was coupled to H-5, H-4 and the methyl group with coupling constants of 7.2, 1.2 and 0.9 Hz, respectively.

The assignment of the ^{13}C spectrum of **5** was straightforward based on the HMQC 2D spectrum [22], and the order agreed with that of 7-methylindole [25]. The quaternary aromatic carbons were assigned based on the assignments of other methyltryptophans, and were simplified because of the 25 ppm range of signals (112.9, 128.7, 121.6 and 137.6 ppm).

Discussion.

Even though there have been several reports on the methyl SCS on indole [24-25,27], ^{1}H nmr spectra of methyltryptophans **2-5** were not previously assigned. Among them, the assignment of **2** is specially important because there are many ergot alkaloids and related compounds bearing this structure [28]. It is interesting that **2** and **5** show CH_3-^{1}H long range couplings up to 6 bonds away, yet do not show detectable 5 bond couplings. Furthermore, it is more interesting that the methyl group of **3** and **4** have long range couplings with only one of the *ortho* protons, adjacent to the ring junction.

After we assigned all ^{13}C resonances of **1-5**, we calculated the methyl SCS for tryptophan. The methyl SCS was determined by subtracting the ^{13}C CS for tryptophan from that of methyltryptophans, and the results are shown in Table 3. The methyl SCS for the benzene ring of indole of methyltryptophans are 9-12 for α-carbons, -2~2 for β-carbons, -0.4-0.5 for γ-carbons, and -2.3~-1.6 for δ-carbons in terms of difference from tryptophan. The changes in the C-3 and C-2 resonances were negligible except in the case of the C-3 resonance of **2**. Like other alkylaromatic compounds, δ-effects are stronger than γ-effects.

We compared methyl SCS for tryptophan with those of indole using the values reported by Fraser and coworkers [25] shown in Figure 4. The correlation is linear to a good approximation with correlation coefficient of 0.98, except in the case of **2**. The ^{1}H CS of one of the β-protons of **2** (3.55 ppm) was different from others (3.31-3.32 ppm), which also suggests the structural changes between **2** and the others. The structural change of **2** was clear when an NOE difference was observed between the methyl group and the α- and one β-hydrogen [22]. This result agreed with the MM2 minimized structure shown in Figure 5 [29]. According to the minimized structure of **2**, the α- and one β-hydrogen were closer to the methyl group than H-5. It was expected that the effect from the structural change of **2** would be concentrated on the carbons around the methyl group and side chain. The large differences of methyl SCS between **2** and 4-methylindole were shown at C-3, C-3a, C-4 and C-5 (Table 4). On the other hand, the fluorine SCS of 4-fluorotryptophan did not show any unusual effects, compared with those of other fluorotryptophans [1]. Thus, the change of **2** may be largely due to steric interaction of the methyl

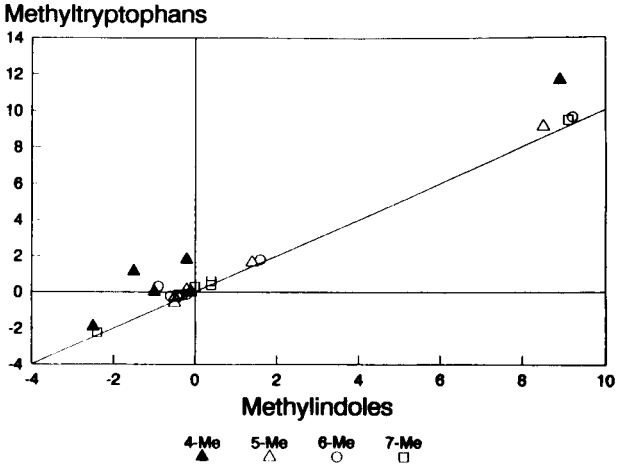


Figure 4. Linear relationship of the methyl SCS between methyltryptophans and methylindoles.

Table 3

The methyl SCS for methyltryptophans

Compound	C-2	C-3	C-3a	C-4	C-5	C-6	C-7	C-7a
1	0.0	1.1	-1.9	11.7	1.7	0.0	-2.0	0.5
3	0.1	-0.6	0.2	-0.4	9.0	1.6	-0.3	-1.6
4	0.3	-0.2	-2.1	-0.3	1.7	9.6	-0.1	0.5
5	-0.2	0.4	-0.4	-2.3	0.2	0.6	9.4	-0.6

Table 4

Methyl SCS of 4-methyl- and 5-methylindole vs 4-methyl- and 5-methyltryptophan

Substitution position	4-methyl indole [a-b]	4-methyl tryptophan	diff.	5-methyl indole [c]	5-methyl tryptophan	diff.
2	-1.0	0.0	-1	0.1	0.1	0.0
3	-1.5	1.1	-2.5	-0.5	-0.6	0.1
3a	-0.07	-1.9	1.83	0.2	0.2	0.0
4	8.9	11.7	-2.8	-0.4	-0.4	0.0
5	-0.2	1.7	-1.9	9.1	9.0	0.1
6	-0.1	0.0	-0.1	1.6	1.6	0.0
7	-2.5	-2.0	-0.5	-0.4	-0.3	-0.1
7a	0.39	0.5	-0.11	-1.7	-1.6	-0.1

[a] The nmr data for C-2, C-3, C-4, C-5, C-6, and C-7 were from the report by Fraser, *et al.* [25]. [b] The nmr data for C-3a and C-7a were from the report by Parker and Roberts [24]. [c] The nmr data were from the report by Morales-Rios, *et al.* [26].

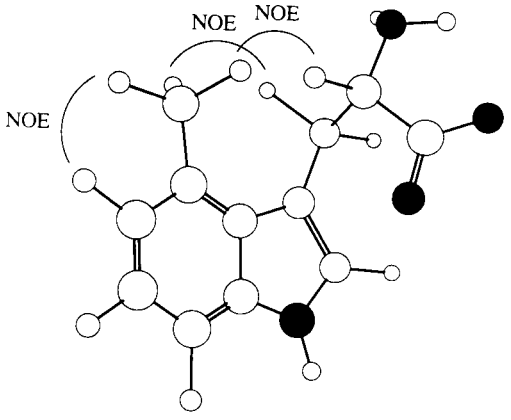


Figure 5. Ball and stick representation of the minimized structure consistent with the NOE data of 4-methyltryptophan.

group with the side chain, as we expected. Whenever the large group SCS of the indoles is applied to the corresponding tryptophans, it should be checked for steric effects in the 4-substituted derivative.

Since there are many alkaloids with alkylated indole skeleta, and they are often hard to obtain in large quantities, the methyl SCS given here will find much more practical applications than previously reported values [30]. The validity of the methyl SCS depends on its utility as that of indole. Even though that was not so impressive in Figure 4, when we compared methyl SCS for **3** with that of 5-methylindole, for which the ^{13}C resonances were assigned very recently [26], the correlation is absolutely linear with a correlation coefficient of 0.999 (Table. 4). This suggests that the methyl SCS of tryptophan could be used as that of indole, except in the case of **2**.

Experimental

The ^{1}H, ^{13}C and NOE difference spectra of **1-5** were recorded on a Bruker AC-300 (300.135 MHz) and AC-250 spectrometers (62.896 MHz), using digital resolution of 0.11 Hz and 0.72 Hz per points, respectively. Solution (20-25 m*M*) in 0.1*N* sodium deuteroxide methanol-d4 were placed in 5mm o.d. sample tubes at room temperature. Chemical shifts were referred to internal TMS.

The HMQC 2D spectra were obtained on a Bruker AC-300 spectrometer using the pulse sequence developed by Bax and coworkers [15]. The fixed delays correspond to a $^{1}J(CH)$ coupling of 150 Hz. Other typical parameters were as follows: spectral width, 1.67 KHz (^{1}H), 4.5 KHz (^{13}C); data matrix, 512 x 512; recycle delay, 2 sec; number of transients, 256; increments of the delay time, 64. Total acquisition time for the data was 11 hours.

The HMBC data were acquired using the pulse sequence of Bax and Summers [16]. The fixed delays correspond to a single long range coupling of 10 Hz. Other typical parameters were as follows: spectral width, 1.67 KHz (^{1}H), 5.8 KHz (^{13}C); data matrix, 512 x 512; recycle delay, 2 sec; number of transients, 256; increments of the delay time, 64. Total acquisition time for the data was 11 hours.

COSY spectra were acquired using a standard pulse sequence employing phase cycling of 45° pulse instead of 90° pulse [23]. Other typical parameters were as follows: spectral width, 1.67 KHz; data matrix, 1K x 512; recycle delay, 2 sec; number of transients, 256; increments of the delay time, 16. Total acquisition time for the data was 3 hours.

Energy minimization was performed on a personal computer using Chem3D Plus molecular modeling software [29]. Compounds **1-5** were purchased from Aldrich Chemical Co. and were used without further purification.

Acknowledgment

We thank Dr. John S. Harwood for technical support and NIH for financial support (GM 42588).

References and Notes

* To whom inquiries should be addressed.

[1] Substituent Effects in Chemical Shifts of Tryptophan Derivatives, Part 2. For Part **1**, M. Lee and R. S. Phillips, *Magn. Reson. Chem.*, accepted.

[2] V. Matarese and D. A. Bernhoffer, *J. Biol. Chem.*, **263**, 14544 (1988).

[3] S. Ali and A. J. Clark, *J. Mol. Biol.*, **199**, 415 (1988).

[4] J. C. Sacchetini, B. Said, H. Shulz, and J. I. Gordon, *J. Biol. Chem.*, **261**, 8218 (1986).

[5] J. M. Beecham and L. Brand, *Annu. Rev. Biochem.*, **54**, 43 (1985).

[6] D. Creed, *Photochem. Photobiol.*, **39**, 537 (1984).

[7] M. S. Morales-Rios and P. Joseph-Nathan, *J. Heterocycl. Chem.*, **30**, 1617 (1986).

[8] C. Winkel, M. W. M. M. Aarts, F. R. van den Heide, E. G. Buitenhuis, and J. Lugtenburg, *Recl. Trav. Chim. Pays-Bas*, **108**, 139 (1989).

[9] E. M. M. van den Berg, F. J. H. M. Jansen, A. T. J. W. de Goede, A. U. Baldew, and J. Lugtenburg, *Recl. Trav. Chim. Pays-Bas*, **109**, 287 (1990).

[10] H. N. Rydon, *J. Chem. Soc.*, 705 (1948).

[11] H. R. Snyder and F. J. Pilgrim, *J. Am. Chem. Soc.*, **70**, 3787 (1948).

[12] U. Hengartener, D. Valentine, Jr., K. K. Johnson, M. E. Larscheid, F. Pigott, F. Scheidl, J. W. Scott, R. C. Sun, J. M. Townsend, and T. H. Williams, *J. Org. Chem.*, **44**, 3741 (1979).

[13] U. Hengartener, A. D. Batcho, J. F. Blount, W. Leimgruber, M. L. Larscheid, and J. W. Scott, *J. Org. Chem.*, **44**, 3748 (1979).

[14] J. Porter, J. Dykert, and J. Rivier, *Int. J. Peptide Protein Res.*, **30**, 13 (1987).

[15] A. Bax, R. H. Griffey, and B. L. Hawkins, *J. Magn. Reson.*, **55**, 301 (1983).

[16] A. Bax and M. F. Summers, *J. Am. Chem. Soc.*, **108**, 2093 (1986).

[17] A. Bax and G. A. Morris, *J. Magn. Reson.*, **42**, 501 (1981).

[18] H. Kessler, C. Griesinger, L. Zarbock, and H. R. Loosli , *J. Mag. Reson.*, **57**, 331 (1984).

[19] B. Bak, C. Dambmann, F. Nicolaisen, and E. J. Pedersen, *J. Mol. Spectrosc.*, **26**, 78 (1968).

[20] F. A. Bovey and G. V. D. Tiers, *J. Am. Chem. Soc.*, **81**, 2870 (1959).

[21] J. S. Cohen, *Biochim. Biophys. Acta.*, **229**, 603 (1971).

[22] M. Lee and R. S. Phillips, unpublished data.

[23] W. P. Aue, E. Barthodi, and R. R. Ernst, *J. Chem. Phys.*, **64**, 2229 (1976).

[24] R. G. Parker and J. D. Roberts, *J. Org. Chem.*, **35**, 996 (1970).

[25] R. T. Fraser, S. Passannanti, and F. Piozzi, *Can. J. Chem.*, **54**, 2915 (1976).

[26] M. S. Morales-Rios, R. E. del Rio, and P. Joseph-Nathan, *Magn. Reson. Chem.*, **26**, 554 (1988).

[27] M. S. Morales-Rios, R. E. del Rio, and P. Joseph-Nathan, *Magn. Reson. Chem.* **27**, 1039 (1989).

[28] I. Ninomiya and T. Kicuchi, in The Alkaloids, Vol **38**, A. Brossi, ed, Academic Press, New York, 1990, pp 1-156.

[29] Chem3D Plus Molecular Modeling Software, Cambridge Scientific Computing, Inc., Cambridge, MA, USA.

[30] Fraser used 0.25 *M* solution in deuteriochloroform. Morales-Rios used 0.83 *M* solution in both deuteriochloroform and DMSO-d_6. On the other hand, Parker used 5 *M* solution in dioxane.