NEW TYPE OF C-GLYCOSIDES OCCURRING IN NATURE

V. K. BHATIA AND T. R. SESHADRI

Department of Chemistry, University of Delhi, Delhi-7

A LARGE number of glycosides, which were earlier isolated from natural sources and studied, were O-linked glycosides; in these, the sugar portion is linked to the aglycone through the oxygen of a phenolic or an alcoholic hydroxyl group. In the study of nucleic acids and the derived nucleosides and subsequently in the study of coenzymes, N-linked glycosides came to be recognised as of common occurrence. Recently glycosides of a novel and stable type have been discovered and these are distinguished by a direct C–C link between the sugar and non-sugar part. Some of these compounds have been known for over a hundred years in the crystalline condition, but it is only during the past ten years that their special nature has been understood. The new type of compounds are called C-glycosyl compounds or C-glycosides. They are rapidly increasing in number and they seem to occur widely. The aglycones involved belong to different types and so far representatives have been found in the groups of anthrones, anthraquinones, flavones, flavonols, flavanones, dihydrochalkones, isoflavones, xanthones and isocoumarins.

The earliest review on this subject was by Hörhammer[1] and a later one by Haynes[2] was more comprehensive. In view of the rapid advances made in this field there is need for frequent reviews. The present article deals with the developments in this field after the review of Haynes. C-Glycosides whose study is fairly complete are briefly described below. They are grouped on the basis of their aglycones for convenience of discussion.

1. ANTHRONE DERIVATIVES

Barbaloin, isobarbaloin, homonataloin and cascarosides A and B and related compounds belong to this group. Barbaloin (I c), one of the earliest known C-glycosides, was isolated in 1851 from the Cape aloes, *A. ferox* and *A. perryi* and Curacao aloes, *A. vera*. It is 9-β-D-glucopyranoside of aloe–emodin anthrone and is accompanied by the α-isomer, isobarbaloin. It has recently been synthesised by condensing aloe-emodin anthrone (I a) and tetra-O-acetyl-α-D-glucopyranosyl bromide (I b). *Aloe speciosa* contains homonataloin, which is 9-D-glucopyranosyl-4, 6-dihydroxy-5-methoxy-2-methyl anthrone. Four aloin-like substances have been isolated from the purgative drug, cascara bark obtained from *Rhamnus purshiana*. Of these, cascarosides A and B seem to be 4- or 5-O-glycosides of barbaloin.[2]

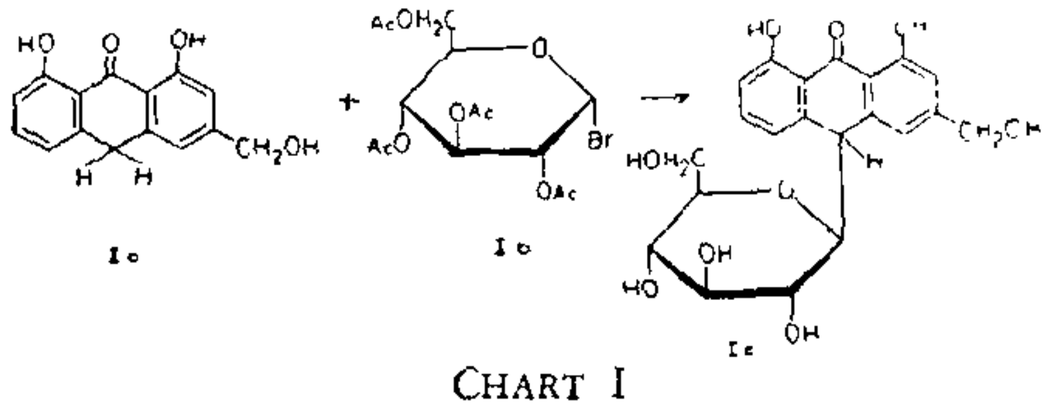
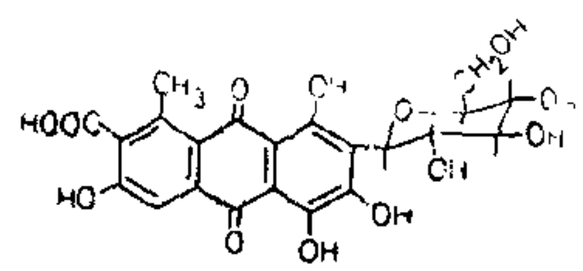


CHART I

2. ANTHRAQUINONE DERIVATIVES

Carminic Acid is the only known anthraquinone C-glycoside and occurs in *Dactylopius coccus* Costa. Its structure has been recently revised to (I d) after the synthesis of degradation products and the NMR spectral studies.[3,4]



I d

CHART II

3. FLAVONE AND FLAVONOL DERIVATIVES

(i) *Vitexin (Orientoside)* was the first C-glycoside to be studied in detail. Recently its isolation has been reported from *Acer palmatum*[5] and *Tamarindus indica*.[6,7] It was originally considered to be apigenin-8-glucofuranoside.[2] The proof that formic acid is a product of periodate oxidation of vitexin by two groups of workers[8,9] has led to the establishment of its correct structure as apigenin-8-β-glucopyranoside (II). This has been supported by NMR spectral studies.[9] Vitexin-4′-rhamnoside occurs in *Crataegus oxyacantha* and *Vitex lucens*.[2] The position of linkage of the rhamnose moiety is unknown.

Isolation of O-D-xylosyl vitexin from *Vitex lucens* and *Citrus sinensis* and p-hydroxybenzoyl vitexin from *Vitex lucens* has been reported.[10] On the basis of NMR spectral studies they have been assigned the constitutions of 2″-O-β-D-xylopyranosyl vitexin (III) and 2″-(p-hydroxy benzoyl) vitexin (IV), the new groups are attached to the concerned alcoholic hydroxyl group in the C-glycosyl part.

(*ii*) *Isovitexin* (*saponaretin, homovitexin*) (V) has also been lately isolated from *Tamarindus indica*.[6,7] It was originally considered to be apigenin-8-C-glucoside with sugar in the open chain form.[2] On the basis of the NMR spectral studies, Horowitz and Gentili[9] have shown that in this compound the β-D-glucopyranose is attached at the 6-position of apigenin. It also occurs as its xyloside and rhamnoside,[2] the position of linkage of these sugars being unknown. Isovitexin-7-methyl ether has been isolated from *Swertia japonica*.[11]

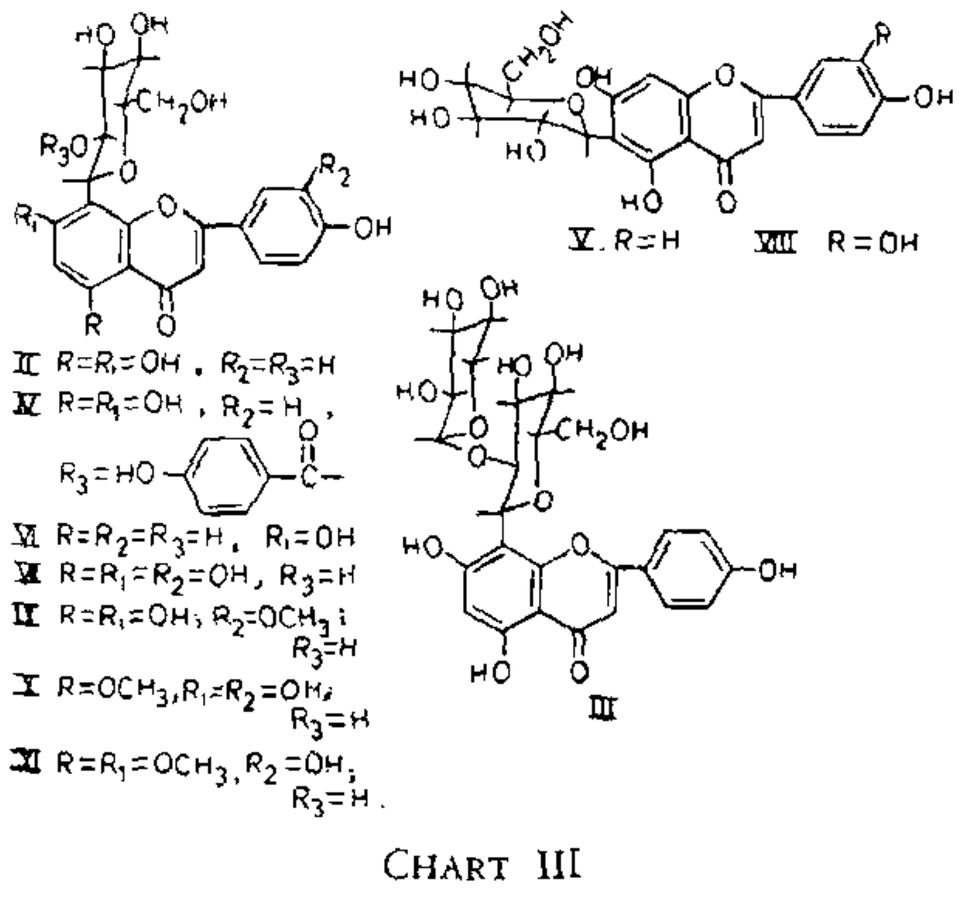


CHART III

(*iii*) *Saponarin and Isosaponarin*.—Saponarin which was earlier isolated from various woods[2] has been recently found to occur in various grasses also.[12] It and isosaponarin are the 7- and 4′-O-glucosides of isovitexin respectively.

(*iv*) *Bayin* occurs in the mature heart-wood of *Castanospermum australe*. Its structure as 5-deoxyvitexin was earlier established by the conversion of 7, 4′-di-O-methylvitexin tetra-acetate into 7, 4′-di-O-methylbayin by the removal of the 5-hydroxyl group.[2] Because of the recent revision of the structure of vitexin, the revised structure of bayin, should be (VI). This has also been supported by the later degradation studies of Eade and co-workers[13] and NMR spectral studies.[14]

(*v*) *Orientin* (*lutexin*) seems to be very widely distributed and has also been isolated from the leaves of *Acer palmatum*,[5] *Helichrysum bracteatum*[15] and *Tamarindus indica*.[6,7] It was earlier considered to be luteolin-8-gluco-furanoside.[2] Koeppen[16,17] found that it consumed 2 moles of periodic acid with the formation of 1 mole of formic acid. The nature of the sugar moiety as glucose was shown by ferric chloride oxidation. On the basis of the above data and NMR spectral studies,[18] orientin has been shown to be luteolin-8-β-glucopyranoside (VII).

Recently Seshadri and co-workers[19-21] have reported the isolation of an epimer of orientin from the leaves of *Parkinsonia aculeata*. This compound (*epi-orientin*) resembles orientin in chemical composition and reactions, chromatographic behaviour, UV and IR spectra. Periodate oxidation of epi-orientin tetramethyl ether leads to the formation of 8-formyl-5, 7, 3′,4′-tetramethoxy flavone. Since orientin is luteolin-8-β-glucopyranoside, epi-orientin should be luteolin-8-α-glucopyranoside.

(*vi*) *Iso-orientin* (*Homo-orientin, Lutonaretin*) has been found to occur in a number of new sources also, *e.g.*, *Helichrysm bracteatum*,[15] *Tamarindus indica*[6,7] and *Swertia japonica*.[11] It was earlier considered to be luteolin-8-C-glucoside with the sugar in open chain form. Periodate and ferric chloride oxidations and NMR spectral studies have led to its structure as luteolin 6-glucopyranoside (VIII). In view of its isomeric relationship, Seshadri and co-workers have suggested that the name homo-orientin may be dropped and iso-orientin[7] used instead.

It also occurs as its xyloside called adonivernoside in *Adonis vernalis*. Iso-orientin 7-O-glucoside (lutonarin) has been isolated from barley leaves along with its 3′-methyl ether.[22] Iso-orientin-7-methyl ether has been found to occur in *Swertia japonica*.[11]

(*vii*) *Scoparin* (*Scoparoside*).—Only a few methyl ethers of C-glycosides have been found to occur in nature and they usually accompany the parent hydroxy compounds. Scoparin, one of the earliest methyl ethers of this type to be isolated, occurs in *Sarothamnus scoparius* and is considered to be orientin-3′-methyl ether based on degradative studies.[2] Because of the recent revision of the structure of orientin, it should have the revised structure (IX).

(*viii*) *Parkinsonin-A and Parkinsonin-B*.—The two compounds have been isolated along with epi-orientin from *Parkinsonia aculeata* by Seshadri and co-workers.[19-21] Both were found to be stable to acidic and enzymatic hydrolysis and on fission with hydriodic acid gave luteolin. Sodium borohydride reduction of the periodate oxidised products of their methyl ethers followed by hydrolysis with acid yielded glycerol. This reaction is given not only by O-glycosides but also by commonly occurring C-glycosides.[6] These methyl ethers consumed 2 moles of periodate with the liberation of 1 mole of formic acid and oxidation with ferric chloride yielded glucose. The above data and

spectral studies lead to the conclusion that parkinsonin-A and parkinsonin-B are 8-β-glucopyranosyl luteolin-5-methyl ether (X) and 8-α-glucopyranosyl luteolin-5, 7-dimethyl ether (XI) respectively.

(*ix*) *Lucenin-1 and Violanthin.*—Recently the existence of di-C-glycosides has been noted. Lucenin-1 was isolated from *Vitex lucens* by Seikel and Mabry.[22] Based on its properties, stability to acidic hydrolysis, its chromatographic behaviour UV and NMR spectral studies it is considered to be luteolin 6, 8-diglucosyl derivative. This source seems to contain many more members of this group.[23] The isolation of another di-C-glycoside, violanthin (6, 8-*diglycoside of apigenin*) has been reported by Hörhamer and co-workers[24] from *Viola tricolor* (garden variety). One of the sugars involved might be rhamnose.

C-glycosyl flavones (6- or 8-C-glycosyl derivatives of apigenin and luteolin) form the dominant flavonoids of the vegetative parts of a wide variety of plants; their occurrence is therefore considered to be useful for taxonomy.[25]

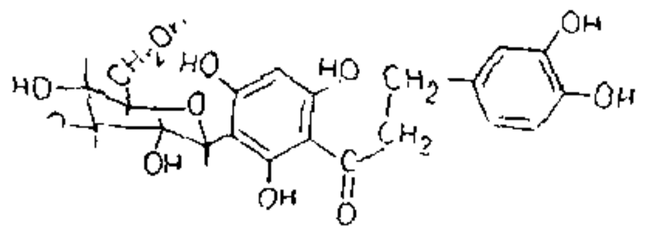
(*x*) *Keyakinin and Keyakinol.*—Keyakinin is considered to be rhamnocitrin-6-C-arabinoside and if correct will be the only flavonol-C-glycoside known.[2] However, according to Harborne,[26] the structural evidence is not conclusive; it is more in favour of its being a flavone than a flavonol. Keyakinol is dihydrokeyakinin.

4. Flavanones

Hemiphloin and Isohemiphloin.—These two compounds were isolated from the kino gum of *Eucalyptus hemiphloia* and the former was originally assigned the structure of naringenin-8-C-glucoside with sugar in the open chain form. Based on its colour reactions, spectral studies and conversion into isovitexin by iodine-oxidation (dehydrogenation), its structure should be revised to naringenin-6-β-glucopyranoside. Isohemiphloin is very similar to hemiphloin and can be obtained from it by prolonged acid treatment and consequently is the position isomer, *i.e.*, 8-C-glucopyranoside.

5. Dihydro Chalkones

Asphalathin.—This compound, isolated from *Asphalathus acuminatus* was earlier considered to be a C-glycosyl derivative of eriodictyol. Detailed study of UV, IR and NMR spectra and periodic acid oxidation lead to the conclusion that it is 3'-C-β-D-glycopyranosyl derivative of 3-hydroxy phloretin and should be represented by (XII).[27a,b] It is the only representative of this group.



XII

Chart IV

6. Isoflavone Derivatives

Puerarin.—Puerarin, daidzein-8-glucopyranoside, is the only known member of this group. It occurs along with its monoxyloside in *Pueraria thumbergiana*,[2] an important Chinese medicament.

7. Xanthone Derivatives

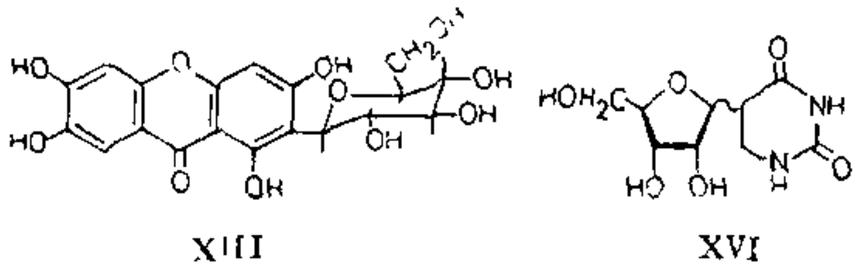
Mangiferin is the only representative of this group and has been known for a long time in the amorphous state but recently has been isolated in the crystalline condition from the leaves, unripe fruits, heart-wood and stem-bark of *Mangifera indica*,[2,28,29] *Anemarrhenæ rhizoma*,[30] and *Hedysarum obscurum*.[1] On the analogy of the structure of euxanthic acid, mangiferin was originally considered to be 7-O-glucoside of 1, 3, 6, 7-tetrahydroxy xanthone. However, since it is resistant to hydrolytic agents it appeared that it belongs to the new class of C-glycosides. That the sugar was D-glucopyranose was shown as follows. Aqueous ferric chloride oxidation of mangiferin gave D-glucose. Mangiferin trimethyl ether consumed 2·2 moles of periodic acid and liberated formic acid and when subjected to periodate oxidation (2 moles) and the resulting dialdehyde treated with sodium borohydride followed by acid hydrolysis gave glycerol.

The position of linkage as 2 was first surmised based on the activity of nuclear positions in tetrahydroxy xanthones and possible biogenesis of C-glycosides and was established by periodate oxidation of the tri-and tetramethyl ethers when the α-hydroxy acetaldehyde of xanthone methyl ethers were obtained as the major fission product. That it was the 2-α-hydroxy acetaldehyde was confirmed by unambiguous synthesis. Osmium tetraoxide oxidation of 1-hydroxy-3, 6, 7-trimethoxy-2-allyl xanthone yielded a diol which underwent periodate fission to yield a xanthone acetaldehyde and further oxidation with one mole of periodate formed the hydroxy acetaldehyde. Quite parallel results have been obtained using mangiferin tetramethyl ether, the resulting hydroxy acetaldehyde synthetically prepared starting from 1, 3, 6, 7-tetramethoxy-2-allyl xanthone. Further evidence has been provided by subjecting the glycoside tetramethyl ether

to oxidation with neutral potassium permanganate whereby the tetramethoxy xanthone-2-carboxylic acid has been obtained.[31,32] Studies of the NMR spectra of mangiferin (XIII) and its derivatives fully support this formulation.[33] In the conversion of mangiferin into euxanthic acid in the animal system, the 3- and 6-phenolic hydroxyls as also the C-C linked glucose unit are obviously removed by a process of reduction involving specific enzymes.

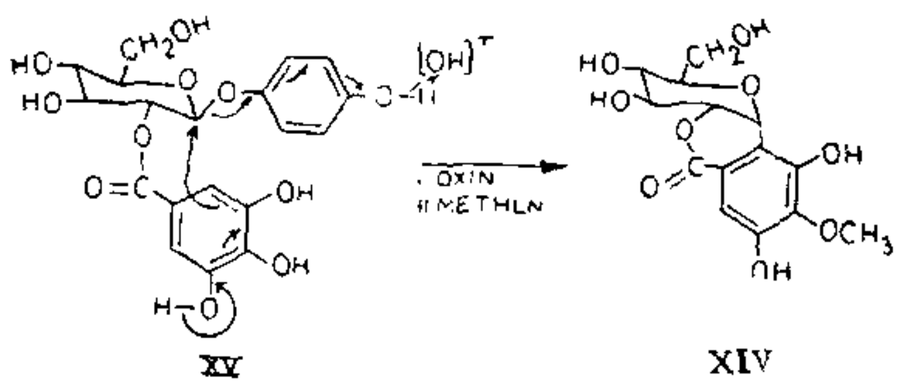
8. Isocoumarin Derivatives

Bergenin (XIV) has been recently found to occur in *Bergenia ciliata, B. strechi*,[34] *Connarus monocarpus*,[35] *Corylopsis spicata* and *Vateria indica*.[36] Synthesis of bergenin was achieved by condensing tetra-O-acetyl-α-D-glucopyranosyl bromide with 4-O-methyl gallic acid in methanol in the presence of sodium methoxide, the O-glucosides formed being subsequently hydrolysed by acid.[2] A possible mode of its biosynthesis from 2-O-galloylarbutin (XV) has been suggested. involving oxidation and methylation as shown in (XV) and (XIV).



XIII XVI

Chart V

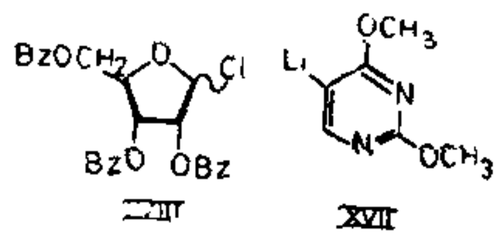


XV XIV

Chart VI

9. Pyrimidine Derivatives

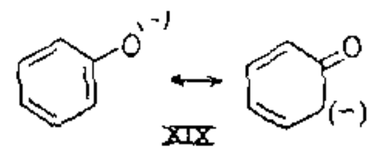
Pseudouridine is a C-glycoside of uracil present in transfer ribonucleic acid and is represented as 5-β-D-ribofuranosyl uracil (XVI). 3, 5-Diribosyl-uracil (5-β-D-ribosyl uridine), occurring in small quantities in certain strains of *E. coli* and of *Penicillium*, is considered to be the intermediate in the formation of pseudouridine. Its synthesis in very low yields was reported by Shapiro and Chambers[38] involving the coupling of 5-lithio-2, 4-dimethoxy-pyrimidine (XVII) with 2, 3, 5-tri-O-benzoyl ribosylchloride (XVIII) followed by hydrolysis with dichloroacetic acid to remove the protecting groups. A modification of this synthesis was to condense 5-lithio-2, 6-ditertiary butoxy pyrimidine with 2, 3, 4, 5-di-O-benzylidine-D-ribose. The removal of the protecting groups and ring closure of the sugar residue was carried out with aqueous methanolic hydrochloric acid.



XVIII XVII

Chart VII

Biopenesis.—The formation of C-glycosyl compounds seems to involve a C-glycosylation process analogous to C-alkylation and can occur either before or after the synthesis of the aglycone; the former route would involve substitution in a poly-β-ketonic precursor. Although no strong argument in favour of one or the other route is available at present, the latter route, *i.e.*, C-glycosylation of the aglycone seems more probable and is supported by the recent synthesis of barbaloin and bergenin. The sugar residues of these substances are attached to highly anionoid centres in the aglycones, most probably as an ultimate step in the biosynthesis. The anion (XIX) of the phenolic aglycone by interaction with a derivative of a 1-phosphorylated sugar would furnish the O- and C-glycoside respectively in a manner exactly analogous to O- and C-alkylation.



XIX

Chart VIII

On analogy with other C-C bond forming reactions in organic synthesis, it is conceivable that the glycosylation might, under physiological conditions, be directed towards an anionoid carbon atom as much as towards an oxygen atom. In this connection it is of biogenetic interest to note the co-occurrence in plants of many C-glycosyl derivatives with their O-glycosides.

The old type O- and N-glycosides undergo rapid hydrolysis in the digestive system of animals and their medicinal properties are thereby considerably affected. They have therefore to be administered by injection. This disability does not exist in the case of C-glycosides. They are generally stable to the action of digestive juices and can function as such in the system. However as mentioned in the case of mangiferin they can be made to undergo reductive cleavage in the animal body but at what site this happens and how, are points to be still investigated.

1. Horhammer, L. and Wagner, H., *Recent Developments in the Chemistry of Natural Phenolic Compounds*, Ed. Ollis, W. D., Pergamon Press, London, 1961, p. 185.
2. Haynes, L. J., *Advances in Carbohydrate Chemistry*, Ed. Wolfrom, M. L., Academic Press, 1963, p. 227.
3. Bhatia, S. N. and Venkataraman, K., *Ind. J. Chem.*, 1965, **3**, 92.
4. Overeem, J. C. and Vanderkuile, G. J. M., *Rec. Trav. Chim.*, 1964, **83**, 1023.
5. Aritoni, *J. Pharm. Soc., Japan*, 1963, **87**, 737.
6. Bhatia, V. K., Gupta, S. R. and Seshadri, T. R., *Curr. Sci.*, 1964, **33**, 581.
7. —, — and —, *Phytochemistry*, 1966, **5**, 177.
8. Rao, C. B. and Venkateswarlu, V., *J. Sci. Ind. Res. India*, 1962, **21 B**, 313.
9. Horowitz, R. M. and Gentili, B., *Chem. Ind.*, 1964, p. 498.
10. — and —, *Ibid*, 1966, p. 625.
11. Komatsu, M. and Tomimori, T., *Tetrahedron Letters*, 1966, p. 1611.
12. Harborne, J. B. and Hall, E., *Phytochemistry*, 1964, **3**, 421.
13. Eade, R. A., Salasoo, I. and Simes, J. J. H., *Aust. J. Chem.*, 1966, **19**, 1717.
14. Hillis, W. E. and Horu, D. H. S., *Ibid.*, 1965, **18**, 531.
15. Rimpler, H., Langhammer, L. and Freuzet, H. J., *Planta Med.*, 1963, **11**, 325.
16. Koeppen, B. H., *Chem. Ind.*, 1962, p. 2145.
17. —, *Z. Naturforsch.*, 1965, **19b** (2), 173.
18. — and Roux, D. G., *Biochem. J.*, 1965, **97**, 444.
19. Bhatia, V. K., Gupta, S. R. and Seshadri, T. R., *Tetrahedron*, 1966, **22**, 1147.
20. —, — and —, *Curr. Sci.*, 1965, **34**, 634.
21. — and Seshadri, T. R., *Phytochemistry* (in Press).
22. Seikel, M. K. and Mabry, T. J., *Tetrahedron Letters*, 1965, p. 1105.
23. —, Chow, J. H. S. and Feldman, L., *Phytochemistry*, 1966, **5**, 439.
24. Horhammer, L., Wagner, H., Rosprim, L., Mabry T. and Rosler, H., *Tetrahedron Letters*, 1965, **22**, 1707.
25. Alston, R. E., Brehm B. G. and McClure, J. W., *Symposium on Newer Trends in Taxonomy*, Nat. Inst. Sci. (India), 1966.
26. Harborne, J. B., *Biochemistry of Phenolic Compounds*, Ed. Harborne, J. B., Academic Press, Inc., London, 1964 p. 146.
27*a*. Koeppen, B. H. and Roux, D. G., *Tetrahedron Letters*, 1965, **39**, 3497.
27*b*. Koeppen, B. H. and Roux, D. G., *Biochem. J.*, 1966. **99**, 604.
28. Ramanathan, J. D. and Seshadri, T. R., *Curr. Sci.*, 1960, **29**, 131.
29. Bhatia, V. K., Ramanathan, J. D. and Seshadri, T. R., *Tetrahedron* (in Press).
30. Morita, N., Shimizu, M. and Fukuta, M. J., *Pharm. Soc., Japan*, 1965, **85**, 374.
31. Haynes, L. J., *Bull. Nat. Inst. Sci.* (*India*), 1965, **31**, 198.
32. Seshadri, T. R., *Ibid.*, 1965, **31**, 200.
33. Haynes, L. J. and Taylor, D. R., *J. Chem Soc.*, 1966, **19**, 1685.
34. Bahl, C. P., Parthasarathy, M. R. and Seshadri, T. R. (Unpublished work).
35. Iyer, S. N., Jain, M. K., Krishnamurti, M. and Seshadri, T. R., *Phytochemistry*, 1964, **3**, 335.
36. Bhrara, S. C. and Seshadri, T. R., *Curr. Sci.*, 1966, **35**, 486.
37. Britton, G. and Haslam, E., *J. Chem. Soc.*, 1965, p. 7312.
38. Shapiro, R. and Chambers, R. W., *J. Amer. Chem. Soc.*, 1961, **83**, 3920.

A NUMERICAL TAXONOMICAL STUDY OF *ARTHROBACTER*

M. F. MULLAKHANBHAI AND J. V. BHAT

Fermentation Technology Laboratory, Indian Institute of Science, Bangalore-12

INCREASED interest in the use of the computers has led us to process the properties recorded of a large collection of freshly isolated cultures of *Arthrobacter* (170 strains) on a IBM 1620 computer. It may be pointed out here that the present-day trend in bacterial systematics is in the direction of recording as many properties of bacteria as possible on their morphology, physiology, nutrition and metabolism. In computing the similarity between two taxonomic entities, numerical taxonomy treats all the taxonomic characters as of equal value and importance. This approach to bacterial systematics has indeed yielded encouraging results in the hands of many investigators.

Among the recently reported investigations on the numerical taxonomy may be mentioned those of Liston,[1] Klinge,[2] Colwell and Liston[3] and Lysenko[4] on the Gram-negative psedoomonads and related forms and those of Focht and Lockhart[5] on the bacterial taxa of both Gram-positive and negative organisms, and of Hill,[6] Pohja and Gyllenberg,[7] Cheeseman and Berridge[8] and Sneath[9] on the Gram-positive genera. The taxonomy of even Streptomycetes has been examined by Silvestri *et al.*[10] and Hill and Silvestri.[11] Numerical taxonomy of the more difficult species as those represented by coryneform bacteria has also been carried out more recently by da Silva and Holt[12] and Harrington.[13]

The present report describes the procedure employed for computer analysis of data recorded for 170 strains of bacteria isolated from glycine enrichments in this laboratory. All the strains were identified as belonging to the genus *Arthrobacter*. A few strains, established by others as those of *Arthrobacter*, were also included in this study. Various characters employed in numerical taxonomy are recorded