Biochemistry (Moscow), Vol. 69, No. 12, 2004, pp. 1379-1385. Translated from Biokhimiya, Vol. 69, No. 12, 2004, pp. 1691-1699.
Original Russian Text Copyright © 2004 by Malygin.

REVIEW

Structure–Chemical Approach to Organization of Information on Metabolic Charts

A. G. Malygin[1,2]

[1]*Bach Institute of Biochemistry, Russian Academy of Sciences, Leninsky pr. 33, 119071 Moscow, Russia; fax: (7-095) 954-2732; E-mail: agmalygin@mail.ru*
[2]*Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, 117997 Moscow, Russia*

Received July 7, 2004

Abstract—A nontraditional approach to construction of metabolic charts is proposed. It is based on the discovery of symmetry in the structure of the network of metabolic reactions. So if compounds and reactions are located on the metabolic chart according to their chemical features, the chart structure will acquire a periodic pattern. The charts thus created have a natural two-dimensional coordinate system of the metabolic network. Points on the X-axis correspond to number of carbon atoms in the carbon skeleton of compounds in columns and points on the Y-axis correspond to number of -COOH groups in compounds filing in series of rows on the charts. As a result this coordinate system sections the field of the charts into rectangular blocks each of which containing compounds with the same numbers of carbon atoms and the same numbers of -COOH groups. The latter substantially improves the charts and makes them a more valid source of metabolic data possessing heuristic properties. The periodicity of the metabolic network structure enables us easily to remember information about biochemical reactions and their products. The charts can also be used as a universal key for any biological database information that is systematically connected with the metabolic information. The charts can be important for medicine and pharmacology because they can help to understand the metabolic processes involved in decomposition of a particular drug or to find the chain of reactions blocked or initiated by administering this drug into a living organism.

Key words: metabolism, metabolic chart, system, structure, symmetry, biochemical reaction, network, information

In this paper, I consider a general approach to organization of material on metabolic charts. It is based on symmetry found in the structure of a network of metabolic reactions. This symmetry is recognized during analysis of similarity and differences in structure–chemical characteristics of metabolites. All the metabolites and reactions coupled to these metabolites are considered as equal participants of a metabolic process, which in turn is sequentially considered from structure–chemical positions. Impairments in symmetry stimulate the search for new metabolites and reactions restoring such symmetry or stimulate the search for convincing external reasons explaining these impairments.

In contrast to the traditional approach to organization of material on metabolic charts based on dogmas on central metabolic pathways, this approach differentiates structure of the metabolic network from its function and considers metabolic pathways as dynamic states of a metabolic network. The unlimited in principle, number of such states excludes complete reflection of metabolic pathways on charts. Attempts of their presentations on the traditional metabolic charts are inconsistent and contradictory. On the contrary, the structural approach is free of such contradictions. For example, the structural approach excludes duplication of inversely directed reactions involved in various metabolic pathways.

On the basis of structural symmetry I have developed a natural rectangular network of coordinates that significantly simplifies the search for required information and facilitates perception of material presented on charts. Impairments of graphic symmetry caused by absence of certain metabolites and reactions on the charts are emphasized by the rectangular coordinate network, thereby making the heuristic properties of these charts more expressed.

GENERAL APPROACH TO ELUCIDATION OF SYMMETRY IN STRUCTURE OF A NETWORK OF METABOLIC REACTIONS AND CREATION OF METABOLIC CHART BASED ON THIS APPROACH

A network of metabolic reactions is a system that implies the existence of its constituents and interrelation-

0006-2979/04/6912-1379 ©2004 MAIK "Nauka / Interperiodica"

ships or connections between these constituents. In the case of a metabolic network, individual metabolites should be considered as elements constituting this network, whereas biochemical reactions responsible for metabolite interconversions represent natural connections.

Like any other systems, the network of metabolic reactions can be characterized by structure and functional states. In the most common form the functional states of this network are determined by concentrations and rates of their mutual inter-transitions in the space of the biosphere, within a framework of separate organisms or their certain parts or fragments (organs, tissues, cells and subcellular compartments). It is clear that the number of various states is countless under these conditions. On the contrary, the structure of the metabolic network is limited; it includes only qualitative description of metabolites and the possibility in principle of occurring of certain biochemical reactions. This implies that in spite of complexity this network is terminate and visible. Elucidation of symmetry in structure of the network of metabolic reactions makes it more available for perception and provides heuristic properties.

The search for symmetry in the structure of the network of the metabolic reactions employs the common definition of an object as property of the object to be consisting of similar parts, which are linked to each other in similar manner.

It is evident that similarity of structure of the network of metabolic reactions can be based on similarity of metabolites and similarity of reactions coupling these metabolites.

Similarity of metabolites is determined by similarity of structure of the carbon skeleton of their molecules and the existence of the same functional groups. Metabolites containing identical functional groups such as -OH, -C=O, -COOH, $-CH_2=CH_2-$, or their certain combinations (e.g., combination of carbonyl and carboxyl groups in 2-keto acids, combination of double bond and carboxyl group at 3,4-unsaturated acids, etc.) are defined as functionally analogous.

Reactions determining similar conversions of some functional groups or their combinations will be similar. For example, oxidation of alcohols to aldehydes, or the reduction of double bonds to single bonds, or decarboxylation of 2-keto acids to aldehydes, etc. are also similar.

In most cases, biochemical reactions meet criterion known in organic chemistry and are defined here as a correspondence principle. According to this principle, compounds with similar structure of molecules participate in similar reactions and the reaction products have also molecules with similar structure.

Based on the definition of symmetry, the search for such symmetry requires classification of metabolites by the existence of similarity in chemical structure. Based on the correspondence principle, functionally analogous metabolites of the same class should be connected to functionally analogous metabolites of the other class by similar reactions.

Figure 1 presents a scheme that shows the principle of the first part of such a search; it consists of two stages.

In the first stage, metabolites are classified by a calculated characteristic (Y1, Y2, Y3). In the case of carboxylic acids the number of carboxyl groups in metabolite molecules was selected as such characteristic. Metabolites with the same number of calculated characteristics were then distributed into rows using qualitative characteristics (a, b, c). In the general case these were characteristics determining functional analogy of metabolites. In the scheme metabolites are shown as rectangles contoured by thin lines; numbers inside these rectangles indicate numbers of calculated characteristics, and the letters of the upright script indicate quantitative characteristics. Regions formed by metabolites characterized by the same number of the calculated characteristics are separated by bold horizontal lines.

At the second stage, the metabolites in the rows were classified by other calculated characteristics (*X1*, *X2*, *X3*). Under real situations, the number of carbon atoms in the carbon skeleton was selected as such a characteristic. Metabolites with the same number of calculated characteristics were in turn separated by qualitative characteristics (*a*, *b*, *c*). These included identity of carbon skeleton branching or hydroxyl group position in the carbon skeleton. The number of calculated characteristics in metabolites is designated in rectangles by inclined numbers and the number of qualitative characteristics is shown by inclined letters. Regions composed by metabolites with the same number of the calculated characteristics are separated by bold vertical lines on the scheme.

Due to intersections of horizontal and vertical separating lines all fields of the scheme are separated into blocks composed of metabolites that are characterized by two numbers of the calculated characteristics of X- and Y-types. These blocks have similar structure and they can be considered as intermediate units of systematization. Separating horizontal and vertical lines form a natural coordinate network.

The second part of this work is demonstrated using a chart of carboxylic acid metabolism as an example (Fig. 2) [1, 2].

The classification of metabolites by their chemical characteristics is the basis of the structure of this chart. For example, classification by number of carboxylic groups is indicated on the chart as three levels of its structure marked with sequential vertical headings on the left side: monocarboxylic, dicarboxylic, and tricarboxylic acids.

Classification on the basis of similarity of other characteristics is shown by rows of functionally related com-

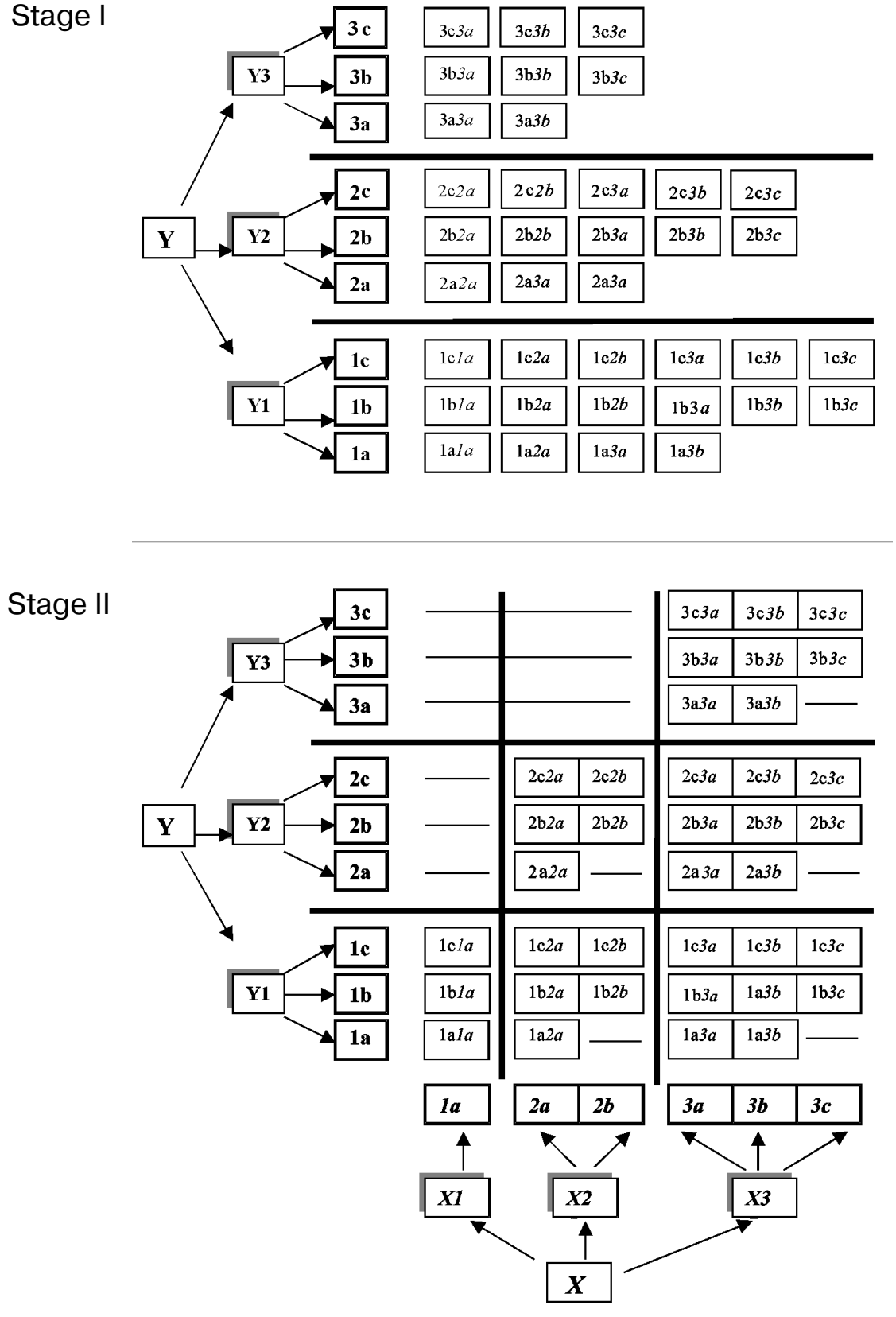

Fig. 1. Scheme of the principle metabolite classification by chemical characteristics.

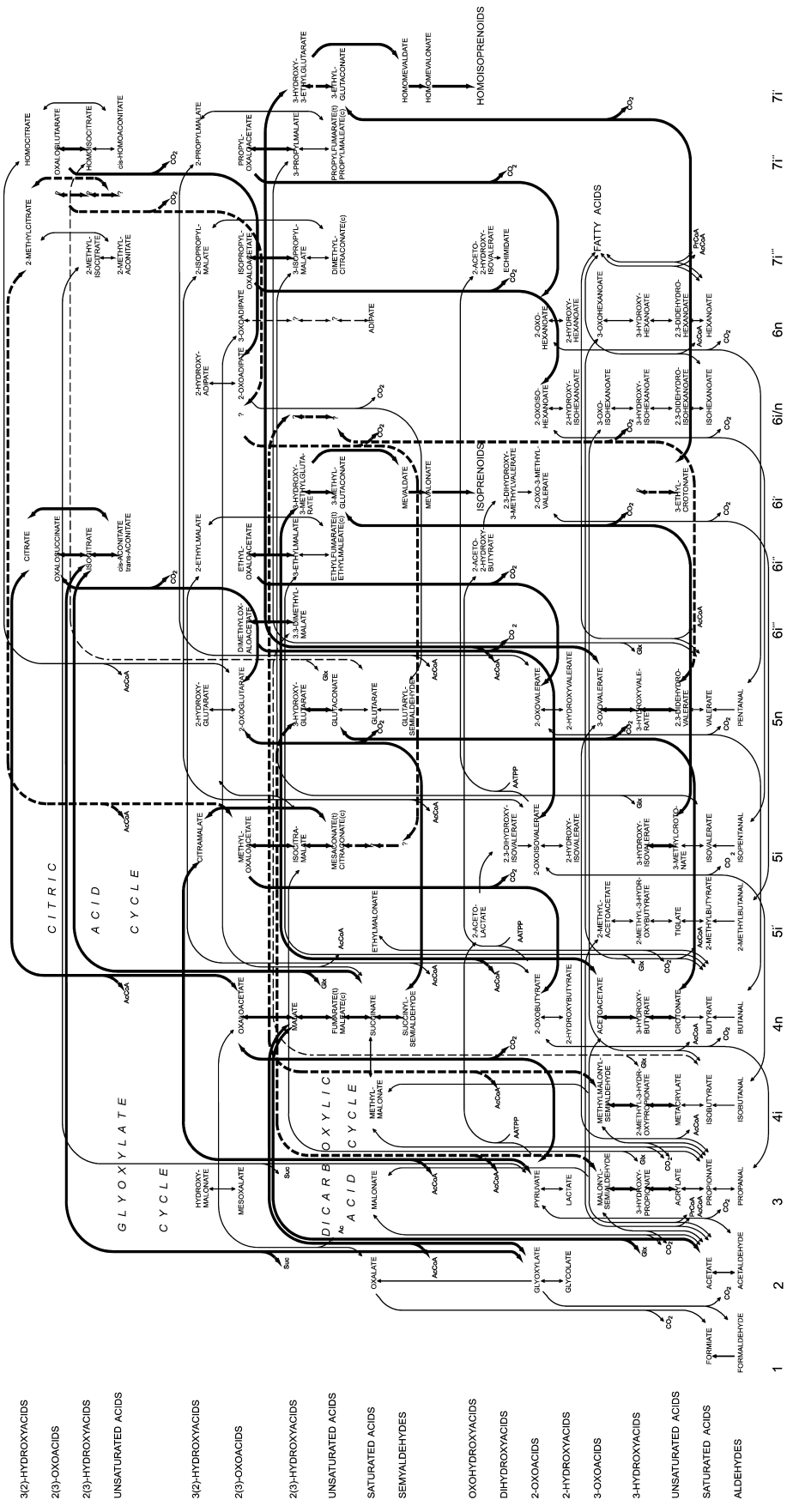

Fig. 2. Structure of the chart of carboxylic acid metabolism.

pounds inside the levels. Names of these rows are shown as horizontal headings on the left side of the chart. It is easy to find that there is a standard sequence of names that is repeated on all three levels. Thus, chart symmetry is manifested as periodicity of its structure on the vertical axis.

Classification by number of carbon atoms in the carbon skeleton of molecules as well as structure of the carbon skeleton is shown in vertical columns, which are marked in the lower part of the chart by corresponding numbers with letter indexes. The letter "n" (normal) shows linearity of the carbon skeleton, and the letter "i" (isomeric) with various numbers of primes shows skeleton isoforms, which are isomeric to the linear one.

This symmetry found within the classification of metabolites determines the symmetry of reactions on this chart; these reactions determine mutual conversions of metabolites. Known reactions are shown with solid arrows, and putative reactions required for maintenance of symmetry are shown with broken arrows. Known reactions represent the major portion of the symmetrical network of reactions in the chart. This means that the correspondence principle can be extended to biochemical reactions.

Coordinate network lines separating blocks on the chart in Fig. 2 are omitted to avoid overloading the chart. However, the position of the blocks can be easily found using coordinate axes and block numeration with horizontal axis indexes and vertical axis letter abbreviations such as: 1-M, 2-M, 2-D, 3-M, 3-D, 4-M, 4-D, 5-M, 5-D, 6-M, 6-D, 6-T, etc.

Using this approach chart structure separates reactions into two large classes. These are reactions linking blocks, reactions changing number of carboxyl groups or sizes of carbohydrate skeleton of reacting molecules, and reactions inside blocks (which do not result in changes of carboxyl group number or changes of size of carbohydrate skeleton of the reacting molecules).

Reactions of metabolite condensation with low molecular weight fragments, decarboxylation reactions, and redox reactions accompanied by appearance/disappearance of carboxyl groups belong to reactions of the first class.

Reactions of the second class include isomerization reactions and reactions accompanied by changes in functional groups which do not result in appearance/disappearance of carboxyl groups.

Charts of metabolism of carbohydrates and nitrogen compounds have been constructed in the same manner [3]. Figure 3 shows a general view of the material organized on the charts. Figure 3 shows vertical lines of the coordinate network common for all three charts; they separate sites of the metabolic network including metabolites with certain carbon number in the skeleton of molecules. These sites are marked with numbers positioned on the horizontal axis.

PRACTICAL APPLICATION OF SYMMETRY RECOGNIZED IN A METABOLIC REACTION NETWORK

The rectangular coordinate network facilitates the use of charts as tutorial or reference material. The regular structure of the charts makes possible their use as a basis for rational systematization of any information systemically related to metabolism [2, 4].

In particular, chart structure is convenient for systematization of inherited and acquired diseases induced by metabolic impairments. For example, the metabolic chart of carboxylic acids accommodates the inborn pathology maple syrup urine disease induced by lack of branched 2-oxo acid decarboxylase and lactacidemia related to pyruvate dehydrogenase deficit in the same row at the level of monocarboxylic acids. 2-Ketoadipic aciduria related to deficit of 2-ketoadipate dehydrogenase activity has similar position at the level of dicarboxylic acids.

Inhibitors of metabolism employed as drugs or venoms may be systemized using the regular structure of the charts. Such systematization would be useful for pharmacology. As an example, one can mention fluoro-substituted citric acid analogs, which can be formed during lethal synthesis from fluoroacetic and corresponding 2-keto acids (by analogy with fluorocitric acid formation from fluoroacetic acid and oxaloacetate in the citric acid cycle). These citric acid analogs will be at a similar position on the chart as corresponding citric acid fluoro-derivative.

Progress in liquid and gas chromatography has provided comprehensive quantitative information about almost all low molecular weight compounds found in organisms or environments. However, due to the complexity of systemic interpretation of data obtained by these methods, they are used very occasionally. Development of metabolic schemes based on new principles significantly facilitates solution of the problem of interpretation of large experimental material obtained during chromatographic analysis of complex mixtures of natural compounds. For example, data on metabolite concentration in tissues and body fluids and their time-dependent changes can be used for chart variants describing the dynamic state of metabolism.

Biochemical reactions occurring in individual organisms in the biosphere are linked into the common network by nutrition chains. Consequently, charts basically describe the metabolism of the whole biosphere. Systemic information on metabolite concentration in aqueous medium, soil, and in the atmosphere presented in the chart structure could function as complex characteristics of environmental state; this can help to elucidate limits of natural fluctuations and determine reasons responsible for exceeding of these limits. Such organization of environmental testing would be more direct and comprehensive characteristics of its ecological state than

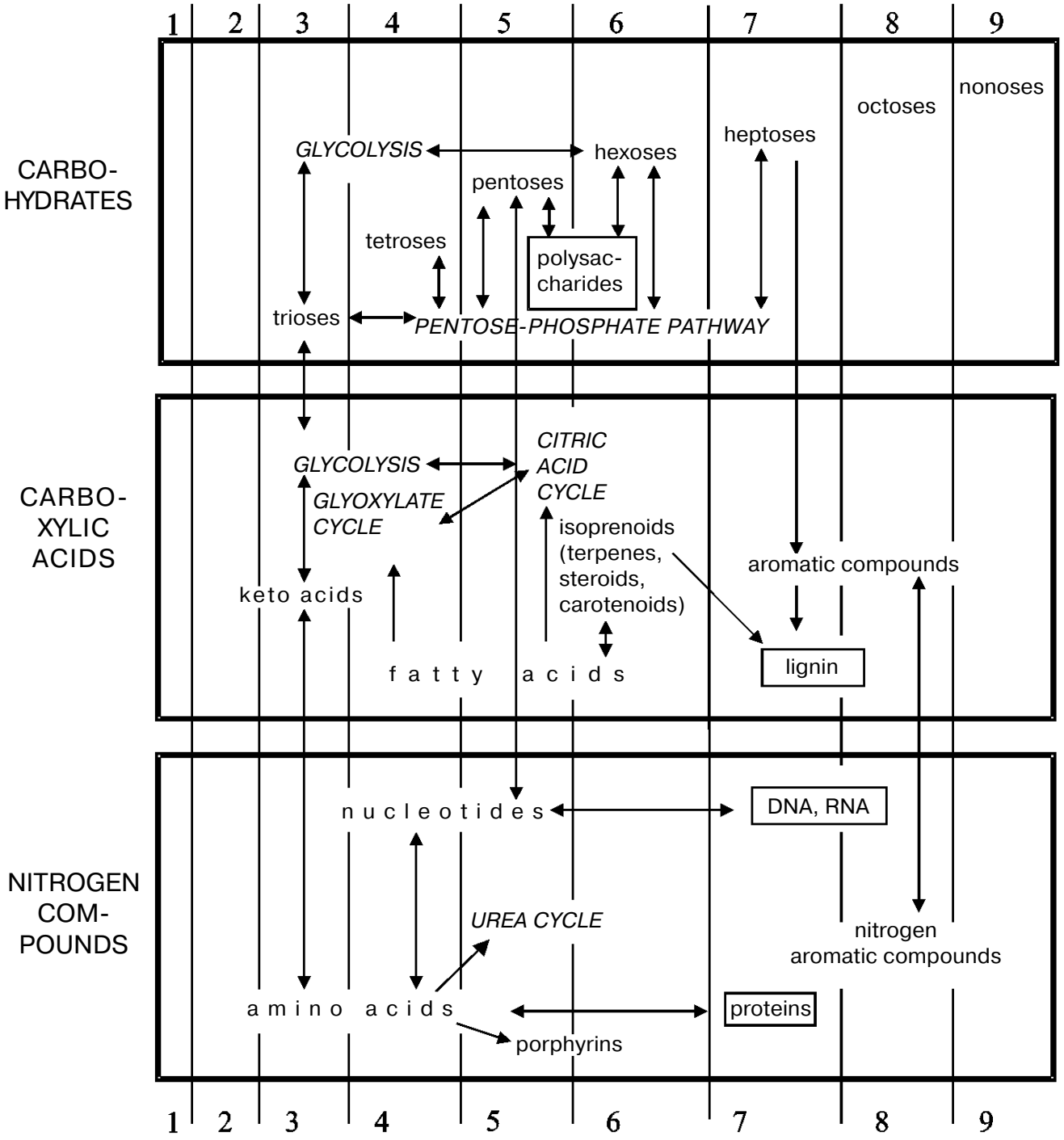

Fig. 3. General view of the material arrangement on the charts of metabolism of carbohydrates, carboxylic acids, and nitrogen compounds.

modern analyses based on evaluation of parameters studied to maximally allowed concentrations.

In experimental studies, charts can be used for planning of artificial metabolic changes by means of gene engineering. In this connection it would be interesting to find elements linking structural symmetry of the metabolic reaction network and functional organization of genomes of various organisms.

Since any symmetry requires its logical completion, structural impairments in symmetry of the metabolic reaction network require special explanations. There are two groups of possible reasons.

The reasons of the first group are related to incomplete knowledge about some sites of the metabolic network. In these case gaps found in the structure of a metabolic network can be used for search and identification in

living organisms of those reactions and metabolites that can restore regular chart structure. In a similar manner, the charts can be used for prognosis (prediction) of any other information systemically related with metabolism. In pharmacology, prognostic properties of the charts may be used for search of natural and artificial regulators of metabolism. For example, fluoro-derivatives of citric acid analogs could be putative metabolic regulators.

Reasons underlying impairments of the second group can consist of the existence of more complex interrelations; this can principally limit manifestations of evident symmetry. For example, such impairments can be caused by incompatibility of some characteristics in structure of one compound. Such variant is arbitrarily demonstrated on the scheme (Fig. 1) as incompatibility of the quantitative characteristics (*X1* and Y2, *X1* and Y3, *X2* and Y3) leading to exclusion of corresponding blocks, and of qualitative characteristics (*2b* and 1a, *2b* and 2a, *3c* and 1a, *3c* and 2a, *3c* and 3a) leading to exclusion of corresponding compounds in each block.

In particular case, from the structure of the carboxylic acid metabolism chart blocks 1-D, 1-T, 2-T, and 3-T are excluded (Fig. 2) because the number of carboxyl groups cannot exceed the number of carbon atoms in the carbon skeleton. Blocks 4-T and 5-T are empty because compounds containing three carboxyl groups in four and five carbon atom skeletons have not been found in nature (due to their instability). The limitation of number of isomers determined by number of carbon atoms in the carbon skeleton represents another demonstrative example of structural limitations. This explains why vertical regions of the chart, which include compounds containing 1, 2, and 3 carbon atoms in the skeleton consist of only one column and number of isomers (and, consequently, number of columns) increases only from the fourth region. To impairments of structural symmetry of the metabolic reaction network (the reasons of which remain to be elucidated) we can attribute the initial asymmetry of the appearance of optical isomers in nature and limitation of amino acid variations in proteins to only twenty structural variants.

Since the information on metabolism is rather large, it is important to summarize and classify it using modern computer technique. However, the problem of its effective use requires deep conceptual analysis. Solution of this problem needs evaluation of possibilities and limitations of computerization applicability and also clear definition of actual problems that could be solved by modern computer techniques. So the concept of use of computer techniques in modern biochemistry implies use of the following functions:

1) provision of effectiveness of reference-search work, which would allow quick construction of a whole information field of the studied problem by fragmented information;

2) natural organization of metabolic information, which suggests convenience of its evaluation and extrapolation on the basis of structural symmetry of the metabolic reaction network;

3) comparative research and summarization of the whole bulk of quantitative experimental information on metabolism obtained by modern methods of chemical analysis.

These systems should serve as diagnostic tools for complex studies of metabolism in living organisms. The periodic structure of the metabolic charts would represent a convenient key for computerized databases of biochemical states in organisms under normal and pathological conditions. Such databases would be very useful in medicine for multi-parameter diagnostics of pathological conditions and for monitoring of the internal medium of the body during medical treatment. Using the structural symmetry of the network of metabolic reactions, facilities of modern chromatographic techniques, and results of analysis of protein content of the living beings by means of two-dimensional electrophoresis in polyacrylamide gel, it is not difficult to develop certain programs of realization of the second and the third functions of computer modeling of metabolism.

This work was supported by the Russian Foundation for Basic Research (grant 04-04-48178a).

REFERENCES

1. Malygin, A. G. (2000) *J. Mol. Med.*, **78**, 66-73.
2. Malygin, A. G. (1999) *Metabolism of Carboxylic Acids (periodical scheme)*, International Educational Program, Moscow.
3. Malygin, A. G. (1992) *Metabolism*, in *Chemical Encyclopedia* [in Russian], Vol. 3, Sovetskaya Entsiklopediya Publisher, Moscow, pp. 610-628.
4. Malygin, A. G. (2002) in *Problems of Biochemistry, Radiation, and Space Biology*, *II Int. Symp. and II Sisakyan Readings*, May 29-June, 2001, Moscow, Dubna, Vol. 2, pp. 105-110.