Glycoconjugate Journal (1994) **11**:169–171

GLYCOPINION MINI-REVIEW

Carbohydrates in cellular recognition: from leucine-zipper to sugar-zipper?

DOROTHE SPILLMANN*

Department of Medical Biochemistry, University of Turku, FIN-20520 Turku, Finland and Friedrich Miescher Institute, CH-4058 Basel, Switzerland

Cells exhibit social behaviour which is crucial for tissue development and maintenance. Recognition and adhesion are key determinants in this social behaviour. What is the basis of cell recognition and adhesion? Is it an all or nothing, hit or miss type of interaction? Or is it rather a multistep process where cell surfaces are tested and primary connections either strengthened or released? Research over the past decades has shown that such a gradual increase in binding strength mediated by multiple binding mechanisms is physiologically significant [1]. How then does this process occur on the molecular level?

Protein–protein interactions can provide tight links by single high affinity binding sites. Moderate affinity between proteins, on the other hand, may be used as a regulative force in cell adhesion phenomena [2]. But is it only adhesive proteins which are capable of establishing specific links between cells? Given that cell surfaces are studded with complex patterns of carbohydrate structures we would like to think that these molecules may also mediate specific interactions. The recognition of specific carbohydrate sequences by lectins has been known for a long time [3], and the importance of such adhesion mechanisms are attracting more and more interest [4]. The lower affinity of such protein–carbohydrate interactions is compensated for by multiple identical binding sites and clustered presentation of carbohydrate ligands. Thus the overall strength of interaction, or avidity, is considerably increased. Further along this conceptual road is the idea of carbohydrate–carbohydrate interactions mediated by specific carbohydrate sequences in an ordered polyvalent array [5].

During the past decade, two different lines of research have lead to a very similar concept, the involvement of polyvalent low affinity carbohydrate–carbohydrate interactions in specific recognition and adhesion, as a first step of a multistep cell-interaction scheme. In marine sponge cell aggregation, polyvalence of low affinity interaction sites has been interpreted as an important determinant in species-specific recognition [6]. Species-specific recognition and aggregation of *Microciona prolifera* cells is mediated by a large proteoglycan-like molecule [7]. Based on work with isolated glycans from this aggregation promoting proteoglycan-like molecule and monoclonal aggregation-inhibiting antibodies the interaction sites were localized in the carbohydrate portion of the molecule [5]. The postulated carbohydrate–carbohydrate interaction between two aggregation molecules provides sufficient binding strength due to a polyvalent arrangement of the interaction sites. The unique arrangement of such sites would furthermore offer the specificity needed to discriminate between similar molecules of different species [6]. This does not exclude other types of cell interaction mechanisms, for instance cellular receptors could work in a lectin-like manner. Secondary, tighter links between cells could be established after cells are sorted and aggregated in this initial, fast step [8].

Characterization of the carbohydrate structures recognized by the antibodies against the sponge adhesion molecule will help us understand sponge cell aggregation and, more generally, principles of cell interactions. Although general mechanisms developed several million years ago might have altered in detail, they may basically still be used by nature in more highly developed organisms. This could be exemplified by studies of mammalian cell systems as described below. Furthermore, we may learn about the evolution of specific carbohydrate structures and their functions as well. A carbohydrate motif, recognized by one of the antibodies inhibiting sponge aggregation, has been shown to be Pyr-4,6-Galβ1-4GlcNAcβ1-3Fuc [9]. This

Abbreviations: G_{M3}, II^3 NeuAc-Lac-Cer; Le^x = Lewis X antigen, Galβ1-4[Fucα1-3]GlcNAcβ1-; SSEA-1, stage specific embryonic antigen.

* To whom correspondence should be addressed: University of Uppsala, Biomedical Centre, Dept of Medical and Physiological Chemistry, 5-751 23 Uppsala, Sweden.

structure is closely related to Siaα2-6Galβ1-4GlcNAcβ1-3Fuc found to be *O*-linked to Ser-61 in the *N*-terminal EGF-domain of human serum factor IX [10]. What makes these structures so interesting for nature to conserve this motif? Does the implication of the first in cell interaction have any implication on the function of the second? We have no idea as yet, but with the advent of modern technology (molecular design, spectroscopy) and improved methods (combination of enzymatic and chemical approaches in both analysis and synthesis) we will be able to create tools for studying such questions.

A multivalent Le^x determinant, identical to the surface SSEA-1 antigen which is highly expressed during the early phase of mouse development, can lead to decompaction of the morula stage mouse embryo while monomeric Le^x or similar control structures do not [11]. This preliminary study has stimulated extensive work on the binding capacity of carbohydrates. For several mammalian cell systems it has been shown that a glycolipid–glycolipid interaction can provide an important recognition and adhesion mechanism (for a review see [12]). The metastatic potential of B16 mouse melanoma cell variants was correlated with their relative degree of G_{M3}-ganglioside expression. The G_{M3}-ganglioside expression was shown to affect the cells' binding capacity to the respective glycolipid-coated surfaces, suggesting a role for such a carbohydrate–carbohydrate interaction in tumour metastasis [13]. In the described model systems – the marine sponge as well as the mammalian cell systems – polyvalence of specific carbohydrate sequences seems to be of crucial important to cell binding. Only by clustering of recognition units is an overall binding strength created, which is high enough to be of value to the cell.

How do carbohydrates interact? Due to their unique features carbohydrate structures offer a multitude of interaction sites. The shape of a carbohydrate chain is determined by the basic structure of the rings, hydrophilic groups, hydrophobic stretches, charged groups, protruding residues and the multiple ways by which the residues are linked together. This massive variation inherent in the composition of carbohydrates distinguishes them from peptides which are strictly defined by the peptide bond motif. Much is known about the forces ruling secondary and tertiary conformations of proteins. The conformations of carbohydrates, though far less characterized, are much more flexible. There seems to be, however, stretches in sugar chains which have a narrow range of optimum conformations [14]. The combination of the two characteristics – the free flexibility of some linkages and the ordered conformation of others – would provide carbohydrates with an excellent recognition testing system. Important recognition sites could be locked in an optimum conformation while their easy presentation would be enabled by a flexible sugar chain.

How are cells kept together by specific carbohydrate chains once the interacting sequences have found each other? It is tempting to imagine that the zipper [15], a simple but efficient way to keep two interacting surfaces together, would also be a suitable mechanism for binding between two complementary carbohydrate chains. For such a model two conditions must be fulfilled: (i) there must be complementary recognition and interaction sites; and (ii) they must be arranged in an ordered, polyvalent array to ensure sufficient binding strength. There are multiple ways to create a repetitive arrangement which could provide the necessary polyvalence. For instance regularly spaced sequences can be located within a linear carbohydrate chain. Alternatively, stubs can protrude from an inert backbone which can be a protein as in mucins, a lipid as with clustered glycolipids, or a carbohydrate backbone as in glycosaminoglycans (Fig. 1). Binding would occur between the surfaces of specific recognition sequences while the scaffolds would not participate in the binding as such. The scaffolds would then determine the overall architecture and thus affect specificity and strength of binding. In the case of freely mobile anchors such as glycolipids, this architecture could be easily changed with a concomitant increase or decrease of binding strength.

The molecular forces active between carbohydrates are the same as between any other type of biological matter. Ionic interactions and hydrogen bonds, hydrophobic

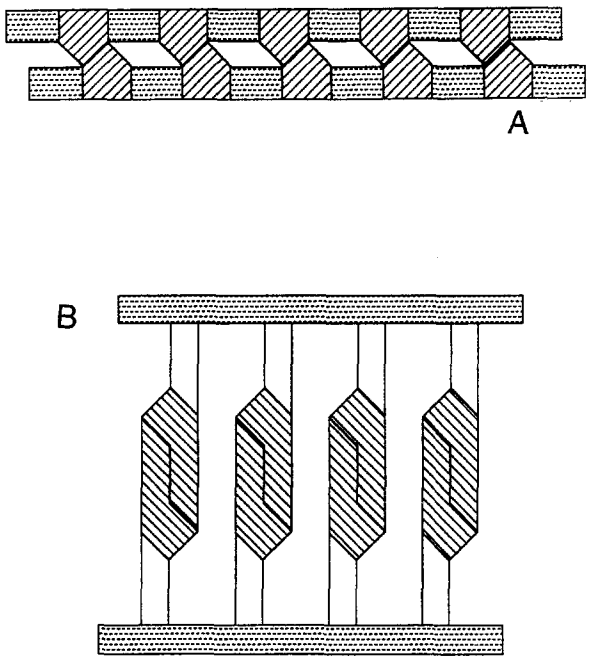

Figure 1. The sugar zipper as a possible model of carbohydrate-carbohydrate interaction sites. In this schematic representation striped areas indicate surfaces involved in recognition and binding, while dotted areas affect the architecture and thereby the overall binding strength and specificity. **A**, all carbohydrate structure; **B**, backbone which can be either protein, carbohydrate or lipid.

interactions and van der Waals forces. Besides hydrogen bonds, ionic interactions do have an important role in carbohydrate interactions. Most carbohydrate chains have a net negative charge, due to the prevalence of diverse acidic groups. Basic groups, on the other hand, are far less common than on peptides. Therefore direct ionic interactions between carbohydrate chains are not likely unless binding is assisted by cations. Such divalent cations have a dual role, both to reduce the effect of repulsion and to increase the strength of binding. This offers another way for fast regulation of binding strength. Changes in the cation concentration can very rapidly change the balance of adhesive versus repulsive forces.

Why then do we know so little about functional interactions of carbohydrates when they offer so many possibilities? Have carbohydrate–carbohydrate interactions been overlooked so far because of the low affinities of the single binding sites? Are there synergistic mechanisms which are used by the cells, whereas the researcher only studies the one with the highest affinity? Or is it the lack of knowledge about carbohydrate structures beyond the most typical structures that sets us so far behind researchers of proteins and nucleic acids? Can we use simplified systems when it is not a question of one binding mechanism but many? What can we learn from these systems? There are more questions than can yet be answered, but above all we have to keep our eyes open for new ideas and concepts. There are enough carbohydrate structures expressed on a single cell in a stage-specific manner for which we still do not have the slightest idea about their *raison d'être*.

Acknowledgement

I wish to thank Drs M. M. Burger, J. Finne, S. Haataja and G. Misevic for critical discussion and continuous support and Dr S. Edwards for correcting the language.

References

1. Hynes RO, Lander AD (1992) *Cell* **68**:303–22.
2. Van der Merwe PA, Brown MH, Davis SJ, Barclay AN (1993) *EMBO J* **12**:4945–54.
3. Barondes SH (1981) *Annu Rev Biochem* **50**:207–31.
4. Lasky LA (1992) *Science* **258**:964–69.
5. Misevic GN, Finne J, Burger MM (1987) *J Biol Chem* **262**:5870–77.
6. Burkart W, Burger MM (1981) *J Supramol Struct Cell Biochem* **16**:179–92.
7. Humphreys T (1963) *Dev Biol* **8**:27–47.
8. Burger MM, Misevic G (1985) in *Cellular and Molecular Control of Direct Cell Interactions* (Marthy HJ, ed) pp. 3–26. New York: Plenum Publishing Corporation.
9. Spillmann D, Hård K, Thomas-Oates J, Vliegenthart, JFG, Misevic G, Burger MM, Finne J (1993) *J Biol Chem* **268**:13378–87.
10. Harris RJ, van Halbeek H, Glushka J, Basa LJ, Ling VT, Smith KJ, Spellmann MW (1993) *Biochemistry* **32**:6539–47.
11. Fenderson BA, Zehavi U, Hakomori S (1984) *J Exp Med* **160**:1591–96.
12. Kojima N (1992) *Trends Glycosci Glycotechn* **4**:491–503.
13. Kojima, N, Hakomori S (1991) *J. Biol Chem* **266**:17552–58.
14. Bush CA (1992) *Curr Opin Struct Biol* **2**:655–60.
15. Landschultz WH, Johnson PF, McKnight SL (1988) *Science* **240**:1759–64.