

An Insulin-Like Modular Basis for the Evolution of Glucose Transporters (GLUT) with Implications for Diabetes

Robert Root-Bernstein

Department of Physiology, 2174 Biomedical and Physical Sciences Building, Michigan State University, East Lansing, MI 48824 U.S.A.

Abstract: Glucose transporters (GLUT) are twelve-transmembrane spanning proteins that contain two pores capable of transporting glucose and dehydroascorbate in and out of cells. The mechanism by which transport is effected is unknown. An evolutionarily-based hypothesis for the mechanism of glucose transport is presented here based on reports that insulin has multiple binding sites for glucose. It is proposed that insulin-like peptides were incorporated as modular elements into transmembrane proteins during evolution, resulting in glucose transporting capacity. Homology searching reveals that all GLUT contain multiple copies of insulin-like regions. These regions map onto a model of GLUT in positions that define the glucose transport cores. This observation provides a mechanism for glucose transport involving the diffusion of glucose from one insulin-like glucose-binding region to another. It also suggests a mechanism by which glucose dysregulation may occur in both type 1 and type 2 diabetes: insulin rapidly self-glycates under hyperglycemic conditions. Insulin-like regions of GLUT may also self-glycate rapidly, thereby interfering with transport of glucose into cells and disabling GLUT sensing of blood glucose levels. All aspects of the hypothesis are experimentally testable.

Introduction

Root-Bernstein and Dillon have suggested that molecular complementarity acts as a selective pressure that directs evolution in particular directions by linking molecular structure and function. In essence, those molecules that bind to one another are more likely to survive degradative processes and so become the materials upon which living systems build. Further, the complexes that such pairs of molecules will form will have emergent properties that their individual components do not, thereby providing a constant source of novelty upon which further complexity can be built. (Root-Bernstein and Dillon, 1997; Hunding et al. 2006) Root-Bernstein recently applied this theory to the elucidation of insulin receptor evolution. Insulin both self-aggregates and binds to glucagon, making insulin self-complementary and complementary to glucagon. (Root-Bernstein and Dobbstein, 1999) Notably, the insulin receptor has multiple insulin-like and glucagon-like sequences located in regions associated with insulin binding, and these regions are highly conserved throughout evolution (Root-Bernstein, 2004; Root-Bernstein, 2005).

Another interesting aspect of insulin behavior is that glucose binds directly to it, suggesting that glucose and insulin were linked functionally very early in evolution. Various physicochemical methods (Anzenbacher and Kalous, 1975; Kalous and Anzenbacher, 1979), x-ray crystallography (Yu and Caspar, 1998), and molecular modeling programs (Zoete, Meuwly and Karplus, 2004) have demonstrated that eight D-glucose binding sites exist on the insulin molecule, two with a binding constant of about 1×10^{-3} M and six with a binding constant of about 6×10^{-2} M. This means that at normal blood glucose of 5×10^{-3} M, the two higher affinity sites are likely to be occupied, while the lower affinity sites are likely to be sensitive only to hyperglycemic conditions. One can hypothesize that prior to the origins of insulin receptors and glucose transport proteins, an insulin-like molecule may have carried glucose into cells itself. In addition, in the presence of glucose, insulin is rapidly and non-enzymatically glycated, which may have a variety of effects to be discussed below.

The binding of glucose to insulin, combined with the molecular complementarity theory of evolution, suggested that perhaps glucose transporters (GLUT) might have evolved from an insulin-like precursor that could also bind glucose. If that were the case, then one would expect to find modular insulin-like regions in the transporter core of GLUT but not at the transporter cores of other sugar transporters. This paper reports that this prediction is correct. It describes how such a structure results

Correspondence: Robert Root-Bernstein, Department of Physiology, 2174 Biomedical and Physical Sciences Building, Michigan State University, East Lansing, MI 48824 U.S.A. Tel: 517-355-6475 ext 1101; Fax: 517-355-5125; Email: rootbern@msu.edu



Copyright in this article, its metadata, and any supplementary data is held by its author or authors. It is published under the Creative Commons Attribution By licence. For further information go to: <http://creativecommons.org/licenses/by/3.0/>.

in functional glucose transport and how glycation of the insulin-like regions in GLUT might result in insulin resistance and inability to transport glucose efficiently.

Methods

Sequence similarities were determined between the human insulin precursor (SwissProt ID P01308) and all human GLUT precursor sequences available on the SWISSPROT database as of November 2006 (GTR1 P11166; GTR2 P11168; GTR3 I1169; GTR4 P14672; GTR5 P22732; GTR6 Q9UGQ3; GTR8 Q9NY64; GTR9 Q9NRM0; GTR10 O95528; GTR11 Q9BYW1; GTR14 Q8TDB8). A further set of similarity searches was carried out in November 2006 involving the sodium/glucose cotransporters 1, 2 and 3 (P13866, P31639, and Q9NY91), using as control sequences the sodium dependent phosphate transporter proteins 1, 2A, 2B, 2C, 3, and 4 (Q14916, Q06495, O95436, Q8N130, O00624, O00476), the sodium dependent phosphate transporter 1 (Q8WUM9), the yeast mannose transporter (P40107), and the arabidopsis mannose transporter (Q941R4).

Similarities were determined using LALIGN (Huang and Miller, 1991). The best 20 similarities were found using default parameters for scoring (penalty for the first residue in a gap, -14; penalty for each additional residue in a gap, -4). Sequences were deemed to be significantly homologous only if they had at least fifty percent identity over a sequence of at least ten amino acids or a raw score (E) of at least 35. Ten amino acid-long sequences were chosen because ten amino acid stretches are approximately the length recognized by T-cell receptors and are therefore of some biological significance (Rudensky et al. 1991) and because in previous studies, it was determined using LALIGN that the probability of a hormone such as insulin or glucagon having five identical amino acids within any stretch of ten amino acids in randomly selected receptors or transporters is six percent ($p = 0.06$) (Root-Bernstein, 2002). A secondary estimation of the significance of the similarities was also used, involving raw scores. Raw scores (E) are used by LALIGN to derive a function that calculates the probability that an alignment with such a score is likely to occur if any peptide of an equivalent length were used to search 10,000 proteins of equivalent length to the receptor or transporter sequence. Raw scores of 35

or above in this study always corresponded to probabilities of less than 1 in 20 that such alignments occur by chance ($p < 0.05$). Raw scores above 50 generally corresponded to probabilities of less than 1 in 100 that such alignments occur by chance ($p < 0.01$). The vast majority of the sequences described in the tables below satisfy both sets of criteria for significance.

It must be emphasized that the probabilities presented above correspond to that of finding a *single* homology between a peptide hormone and a receptor satisfying one or the other criteria. The probability that *more than one* such significant homology would occur within a single receptor or transporter is a multiplicative function of the number of significant sequences found. The probability of finding four such similarities within a single transporter sequence is approximately $(0.05)^4$ or about 6×10^{-6} and the probability of finding seven such significant similarities is about 10^{-11} . All of the GLUT transporters have between four and twelve significant homologies so that the probability that the results presented here have occurred by chance are extremely small.

The sequences of GLUT transporters that had high homology with insulin were mapped onto the model of the GLUT transporter developed by Zuniga et al. (Zuniga et al. 2001) using the amino acid identifiers in their model.

Results

The results of the homology study are displayed in Tables 1 through 14. As can be seen by inspection, every human GLUT sequence currently available on SwissProt has multiple homologies with insulin satisfying the criteria described above (Tables 1–11). Particularly notable are GLUT 1, GLUT 4, GLUT 6, GLUT 10, and GLUT 11. GLUT 1 has eight significant homologies; GLUT 4 has ten; GLUT 6 and 10 have seven, and GLUT 11 has eleven. Most strikingly, GLUT 11 has one homology that represents virtually the entire unprocessed insulin chain. Many significant similarities were also found between the sodium/glucose cotransporter proteins and insulin (Tables 12–14).

No significant similarities at all were found between insulin and the sodium dependent phosphate transport proteins, the sodium dependent phosphate transporter, or mannose transporters (data not shown, since there is nothing to show). These results are in accord with our previous

Table 1. GLUT 1-Insulin similarities.

INS 81–87	ALE <u>GSLQ</u>	
GLUT 1 17–25	GGAVL <u>GSLQ</u>	34E
INS 70–78	GGPGAG <u>S</u> LQ	
GLUT 1 56–93	I LPT–T LTTLW SLS – VA I FSVGG M IGSFSVGLFV NRFGRR	
INSULIN 7–46	LLPLLALLALWGPDPAAAFV NQH LCGSHLVEALYLVCGER	
INS 84–92	GS LQ <u>RGIV</u>	
GLUT 1 137–165	TGFVPMYVGEVSPTAFRGALGTLH <u>QLGIV</u>	37E
INS 67–86	REAEDLQV <u>GSLQ</u> PLALEGSL	
GLUT 1 181–190	GNKDLWPLLL	27E
INS 73–82	GAGS <u>LQPLAL</u>	
GLUT 1 278–286	LQLSQ – QLSG	21E
INS 61–70	LQVGQVELGG	
GLUT 1 310–329	AT I GSGIVNTAFT–VVSLFVVE	26E
INS 90–1XX	<u>SIQKRGIVEQCCTSICSLYQLE</u>	
GLUT 1 353–394	WMSYLSIVAIF GFVAFFE V GPGPI PWF IVA EL	38E
INS 4–30	WMRLLPLLALLALW – – – – – GPDPAAAF VNQHL	
GLUT 1 390–406	IVAELFAQ GPRPAAIA V	40E
INS 10–26	LLALLALWGPDPAAAFV	
GLUT 1 449–471	YFKVPETKGRTFDEI ASGF RQGG	36E
INS 48–70	FFY TPKTRREAEDLQVGQVELGG	

observation that insulin-like regions are extremely rare in randomly selected receptors and enzymes (Root-Bernstein, 2002) It is also interesting to note that mannose transporters lack significant homologies to the QLSQQLS, and QLG sequences that have been identified through amino acid substitution and modeling as probable glucose recognition regions in GLUT (Seatter et al. 1998; Olsowski et al. 2000, Salas-Burgos, 2004).

On the other hand, insulin itself contains several regions that have the QLS and QLG motifs associated with insulin binding—although these are sometimes in the reversed order—and these motifs

occur in regions with significant homologies to GLUT transport cores (see underlines in Tables 1–11). Thus, the insulin-like homologies reported here correspond to putative glucose binding regions identified by previous investigators of GLUT structure (Seatter et al. 1998; Olsowski et al. 2000, Salas-Burgos, 2004; Hruz and Mueckler, 2001; Zuniga et al. 2001).

In order to demonstrate the location of these insulin-like regions in the GLUT structure, the homologies listed in Tables 1–11 (the GLUT sequences) have been mapped onto Zuniga et al.'s model of GLUT transport cores (Zuniga et al. 2001).

Table 2. GLUT 2-Insulin similarities.

GLUT 2 176–188	GEIA PTALRGAL	41E
INS 75–86	GSLQPLALE GSL	
GLUT 2 217–230	LWHI LLGLSGVRAI	30E
INS 3–16	LWMRLLP LLALLAL	
GLUT 2 174–196	MYIGE IAPTALR GALGTFHQ LA I	28E
INS 61–82	LQVGQVELGGPGA–GSLQPLAL	
GLUT 2 406–417	LFVSFF E I GPG P	33E
INS 61–72	LQVGQVELGGGP	
GLUT 2 430–435	GPRPAA	36E
INS 18–23	GPDPAA	

To simplify the visualization of the placement of the various homologies, only those shared by GLUT 1 and GLUT 4 (arguably two of the most important of the GLUTs) were used; these were reformatted (Table 15) to display their commonalities; and these common areas were then mapped onto Zuniga's model of the transport cores by means of color coded dots (Figure 1). Figure 1

demonstrates that many of the insulin-like regions that are common among the GLUTs can be assigned to regions of the GLUTs thought to define the transporter core itself. Notably, these insulin-like regions are found to be in the transporter cores of all models of GLUT (e.g. Seatter et al. 1998; Olsowski et al. 2000, Salas-Burgos, 2004; Hruz and Mueckler, 2001). Moreover, all of the amino

Table 3. GLUT 3-Insulin similarities.

GLUT 3 61–118	SLWSLS VAI FSVGGM I GSFSVGLFVNRFGRNRSM LI VNLLAVTGG–C FMGLCKVAKSVEM	43E
INS 2–61	ALWMRLLP LLALLALWGPDPAAAFVNQH–LCGSH LVEALYLVCGERGFFYTP KTRREAE D	
GLUT 3 143–154	GE ISPTALRGAF	38E
INS 75–86	GSLQPLALEGSL	
GLUT 3 155–163	GTLNQLGIV	32E
INS 84–92	GSLQKRGIV	
GLUT 3 380–406	GPGPIPW FIVA ELF SQ GPRPAAMAVAG	36E
INS 18–43	GPDPAAAFVNQHLCGSH LVEALYL VCG	
GLUT 3 388–404	IVAELFS Q GPRPAAMAV	40E
INS 10–26	LLALLALWGPDPAAAFV	
GLUT 3 447–461	FFKVPETRGRTFEDI	42E
INS 48–61	FFYTPKTR– REAEDL	

Table 4. GLUT 4-Insulin similarities.

GLUT 4 31–37	AVLGS <u>LQ</u>	28E
INS 81–87	: ALEGS <u>LQ</u>	
INSULIN 7–46	LLPLLALLALWGPDPAAAFV NQH LCGSHLVEALYLVCG ER	
GLUT 4 90–109	: : : : : : : \ : : : : : IPPG–TLTTI <u>WALS</u> – VA I FSVGG M I SSF L IG I ISQWLGRK	
GLUT 4 124–147	SLMGLANAAASYEML ILGRFL IGA	30E
INS 15–38	: : : : : ALWGPDPAAAFVNQHLCGSHLVEA	
GLUT 4 158–180	MYVGE IAPTHLRGALGTLN <u>QLAI</u>	30E
INS 61–82	: : : : : : LQ VGQVELGGPGA–GS <u>LQPLAL</u>	
INS 2–16	ALWMRLLPLLALLAL	
GLUT 4 197–214	: : GTASLWPLLL G LTVLPAL	38E
INS 73–82	: : GAGSLQPLAL	
GLUT 4 204–238	LL LG LTVLPALLQV – LLP FCPEsprYLY I IQNLEG	35E
INS 1–32	: : : : : : : : : : : : MALWMRLLPLLALLAL—WGPDPAAAFVNQHLCG	
GLUT 4 273–286	LSLLQL LGSR THR	28E
INSULIN 79–92	: : : LQPLALEGS <u>LQKR</u>	
GLUT 4 294–302	<u>LQLSQ</u> – QLSG	21E
INS 61–70	: : : : <u>LQVGQVELGG</u>	
GLUT 4 307–330	FYYSTS I FETAGV <u>GQPA</u> YATIGAG	33E
INSULIN 47–71	: : \ : : : : : : FFYTPKTRREAED <u>LQVGQVELGGG</u>	
GLUT 4 350–362	RTLHLLGLAGMC G	28E
INS 6–18	: : : RLLP LLALLA LWG	
INS 18–44	GPDPAAFVNQHLCGSHLVEALY LVCG	
GLUT 4 398–424	: : : : : : GPGP I PWF IVAEL F SQGPRPAAMAVAG	40E
INS 10–26	: : LLALLALWGPDPAAAFV	

acids thus far identified as contributing to glucose recognition and binding through mutagenesis studies fall into these insulin-like regions or immediately adjacent to them, including GLUT 1 residues W65, S66, T137, V140, QLS 161–163, V165, QLS 279–285, T310, N317, T321, R333, P387,

and W388 (Seatter et al. 1998; Olsowski et al. 2000; Salas-Burgos, 2004; Hruz and Mueckler, 2001; Zuniga et al. 2001).

These data therefore show that the transport cores of GLUT, as defined by current experimental and modeling data, are composed of repeating

Table 5. GLUT 5-Insulin similarities.

GLUT 5 12–22	RLTLVLALATL	26E
	: :	
INS 6–16	RLLPLLALL AL	
GLUT 5 85–100	GSLLVGPLVNFGRKG	30E
	: :	
INS 32–47	GSHLVEALYLVCGERG	
GLUT 5 151–172	GELAPKNLRGAL – –GVPQEI TV	38E
	: : : : :	
INS 75–99	GSLQPLALEG SLQKRG I VE QCCTSI	
GLUT 5 214–223	<u>ALQ</u> LL LL PFF	31E
	: :	
INS 2–11	ALWMRLLPLL	

subunits, or modules, based on an insulin-like sequence. It is further reasonable to propose that since insulin has multiple binding sites for glucose, these insulin-like modular regions are the basis for GLUT binding and transport of glucose.

Discussion

The fact that glucose binds to insulin and that insulin-like regions make up the transport cores of

GLUT proteins suggests an obvious mechanism by which glucose can be transported through these proteins. Glucose would be attracted from extracellular plasma to the glucose-binding sites on the insulin-like regions in the transport core. Simple diffusion from one site to another within the transport core would carry the glucose from the extracellular to the intracellular side of the transporter as long as the concentration of glucose was higher outside the cell than inside. Because the proposed

Table 6. GLUT 6-Insulin similarities.

GLUT 6 91–109	GAAAGGLSAMILNDLLGRK	36E
	: : : : : :	
INS 71–89	GPGAGS LQPLALEGSLQKR	
GLUT 6 167–182	GALGATPQLM AVFGSL	28E
	: : : : :	
INS 71–86	GPGAGS LQPLALEG SL	
GLUT 6 278–288	VAL L MRLQQL	
	:	
INS 1–11	MALWMRLLPL L	
GLUT 6 308–66	LLP PKDDAAI VGAVRLLSVLIA ALTMDLAGRKVLLFVSAAIMFAANLTLGLYI HFGPRP	38E
	: : : : : : : : : : : : :	
INS 16–72	LWGPDPAAAFVNQHLCGSHLVEALY L- VCGERGFFYTPKTRREAEDLQVG-QVELGGGP	
INS 2–14	ALWMRLLPLLAL L	
	: : : :	
GLUT 391–403	AGYLTLV PLLATMLFIMGYAVG WGP	38E
	: : : : : :	
INS 3–19	LWMRLLPLLAL LALWGP	
GLUT 6 481–497	CCVPETKGRSLEQIES F	29E
	: : : :	
INS 92–108	CC- - - TSIC SLYQLIN Y	

Table 7. GLUT 8-Insulin similarities.

GLUT 8 24–37	RRVFLAAFAAAL GP	27E
	: :	
INS 6–19	RLLP LLALLALWGP	
GLUT 8 122–133	MLLGG RLLTGLA	26E
INS 1–12	MALWMRLLPL LA	
GLUT 8 170–177	LLAYLAGW	35E
INS 10–17	LLALLALW	
INS 6–20		RLLP LL ALLALWGPD
		: :
GLUT 8 194–236	LMLLLMCFMPETPRF LLTQ H RRQEAMAALRFLWGSEQGWEDPP	42E
	: : : : : \ \ : : :	
INS 10–52	LLALLAL WGPDPAAAFVNQH C G SH LVEAL YLVCGERG F FYTP	
GLUT 8 345–351	GGPGNSS	38E
	:	
INS 70–76	GGPGAGS	

mechanism is passive, diffusion could also occur from the inside of the cell out if the concentration of glucose were higher within the cell than outside it. If the binding affinities were essentially equal from one site within the transport core to the next, then transport would occur with equal velocity in either direction. If the bindings progress either toward greater affinity from outside in, or vice versa, then the velocity of diffusion will differ depending on whether the glucose is diffusing into or out of the cell. The difference in rate will depend on the difference in affinities across the transporter.

Differential affinities might be of value in regulating rates of glucose flow.

The fact that multiple insulin-like regions occur within the transport core should not come as a great surprise in terms of the way in which most proteins are thought to have evolved. An ever-increasing body of literature suggests that large proteins are often conglomerates of repeating subunits or modules (e.g. Patthy, 1999; Liu and Rost, 2004; Bornberg-Bauer et al. 2005). Thus, the structure proposed here for GLUT proteins may be seen as just one of many known modular protein structures.

Table 8. GLUT 9-Insulin similarities.

GLUT 9 156–166	ACS LQAGAFEM	29E
	: : :	
INS 58–68	AE DLQVQVEL	
GLUT 9 299–308	I RLVSVLELL	29E
	: : :	
INS 5–15	MRLPLLALL	
GLUT 9 310–339	APYVRWQVVTIVITMACYQLCGLNAIWFYT	38E
	: : : : : : :	
INS 23–51	AAFVNQ HLCGSHLVEAL YLVCGERG-F FYT	
GLUT 9 321–340	IVTMACYQLCGLNAIWFYTN	45E
	: : : : : :	
INS 91–110	IVE QCCT SIC SLYQLENYCN	

Table 9. GLUT 10-Insulin similarities.

GLUT 10 7–18	VLPLCASVSLG	34E
INS 7–18	LPLLALLALWG	
GLUT 10 30–42	SGALLPLQLDFGL	35E
INS 74–86	AGSLQPLALEGSL	
GLUT 10 87–99	AGSLT–LGLAGSL	35E
GLUT 10 50–78	LVGSLLLGAL – LASLVGGFLIDCYGRKQA	34E
INS 30–58	LCGSHLVEALYLVCGERGFFYTPKTRREA	
GLUT 10 76–84	QAILGSNLV	29E
INS 28–36	QHLCGSHLV	
GLUT 10 271–304	LASVGLGAVKVAATLTAMGLVDRAGRALLLAGC	35E
INS 63–96	VGQVELGGPGAG SLQP LALEGSLQ KRGIVEQCC	
GLUT 10 337–368	GQTGLP GDSGLLQD SSLP PIPRT NEDQREP IL	30E
INS 64–92	GQVE LGGGPGA – – – GSLQPLALEGSLQKRGIV	

What is unusual about this example of modularity is that it may have been based on complementarity between a pair of simple molecules (insulin and glucose) that dictated selection for the specific utility and function of the modules during the evolutionary process itself (Root-Bernstein and Dillon, 1997). Since this is also the case in the insulin receptor, one may speculate that such complementarity will be found in other receptor and transporter proteins as well.

The hypothesis that GLUT evolved from the stringing together of insulin-like modules capable of binding glucose makes a number of novel, experimentally testable predictions. Synthetic peptides derived from the regions of insulin-like homology listed in the Tables above should bind glucose with approximately the affinity previously found for insulin-glucose binding sites. For reasons described above regarding the rate of flow of glucose through the transporter, it will be interesting and important to determine whether all of the binding sites are more or less similar in their glucose affinity, or whether they increase in one direction or the other.

A second prediction of the hypothesis is that these insulin-like regions of GLUT should glycate

rapidly (hours to days) in hyperglycemic conditions. It has been shown by several groups of investigators that insulin in the presence of either unusually high concentrations of glucose, or normal glucose concentrations for longer periods of time, will auto-glycate. (Farah et al. 2005; McKillop et al. 2003; Abdel-Wahab et al. 2002; Boyd et al. 2000; O’Harte et al. 2000; Abdel-Wahab et al. 1997a, b; O’Harte et al. 1996; Dollhofer and Wieland, 1979). This auto-glycation occurs on both N- and O-groups of amino acids; is non-enzymatic; and its effect on insulin is to decrease its activity. In contrast to the very slow (weeks to months) non-enzymatic glycation of most proteins (Schalkwijk, Stehouwer and van Hinsbergh, 2004; Vlassara and Palace, 2002; Baynes, 2001), the glycation of insulin is very fast, beginning within hours and reaching a maximum within 24 to 36 hours in hyperglycemic conditions. (Farah et al. 2005; McKillop et al. 2003; Abdel-Wahab et al. 2002; Boyd et al. 2000; O’Harte et al. 2000; Abdel-Wahab et al. 1997a, b; O’Harte et al. 1996; Dollhofer and Wieland, 1979) One would expect the same rapid glycation to occur on the insulin-like regions of GLUT. This prediction can

Table 10. GLUT 11-Insulin similarities.

GLUT 11 64–158	LLMWSL IVS LYPLGGL F GALLAGPLAITLGRKKSLLVNNI FVVSAA IL FGFSRKAGS : : : : : : : : : : : : : : : : :	
INS 1–92	MALWMRLLPLLALLALWGPDPAAAF–VNQHLCGSHLVEALYLVCGERGFFYTPKTRR F –EMIMLGRLLVGV NAGVSMNIQPMYLSSESAPKELRGAV 52E : : : : : : : : : : : : : : :	
GLUT 11 70–76	IVSLYPL 26E 	
INS 99–105	ICSLYQL	
GLUT 11 72–83	SLYPLGG LF GAL 26E : :	
INS 76–86	SLQPLA – LE GSL	
GLUT 11 69–100	L I VS LYPLGGL F GALLAGPLAITL GRKKSLLV 42E : : : : :	
INS 61–92	LQVGQVE LGGGP GAGSLQPLALEGSLQKRGIV 	
GLUT 11 176–190	VGLRELL GGP QAWPL 45E	
GLUT 11 198–226	PGALQLASLPLLPESP RYLL I DCGDTEAC 33E : : : :	
INS 69–100	GGGPGAGSLQPLALEGSLQKRGIVEQCC TSI C : :	
GLUT 11 235–249	GSGDLAGEL EELEEE 36E	
GLUT 11 286–294	LCGNDSVYA 31E :	
INS 30–38	LCGSHLVE A	
GLUT 11 383–397	FGI GP–AGVTGILATE 29E :	
INS 68–83	LGGGPGAGSLQPLALE	
GLUT 11 386–412	GPAGVTGI LATELFDQMARPAACMVCG 37E : : : : : : :	
INS 18–44	GPDPAAAFVNQHLCGSHLVEALY LVCG	
GLUT 11 446–460	VCGAIYT GLFL –PETK 27E : : :	
INS 42–55	VCGE – –RGFFYT PKTR	

be tested both *in vitro* using insulin-like peptide regions derived from GLUT, or under *in vivo* conditions, in which the GLUT are isolated and analyzed for the presence of glycation products within specific regions of the transport core. If glycation is present under *in vivo* conditions, it would further be predicted that glucose transport would decrease inversely to increasing GLUT glycation.

A third test of the hypothesis stems from the fact that GLUT also transport dehydroascorbate

(DHA) (reviewed in Root-Bernstein, Busik and Henry, 2002). It can be predicted, therefore, that insulin and the insulin-like regions of GLUT should bind DHA with affinities approximating those which they bind glucose. Moreover, DHA binding may significantly retard glycation, which can be tested in both the *in vitro* and *in vivo* conditions just specified above. This prediction may be of clinical significance as well (see below).

Table 11. GLUT 14-Insulin similarities.

GLUT 14 9–30	VSG IGGFLVSLTSRMKPHTLAV	33E
INS 42–63	VCGERGGFFYTPKTRREAEDLQV	
GLUT 14 167–178	GEI SPTALRGAF	38E
INS 75–86	GSLQPLALEGSL	
GLUT 14 398–407	VACFE I GPGP	33E
INS 63–72	VGQVELGGGP	
INS 10–26	LLALLALWGPDPAAAFV	
GLUT 14 404–430	GPGP I PWFI VAELFSQGPRPAAM AVAG	40E
INS 18–44	GPDPAAAFVNQHLCGSHLVEALYLVCG	
GLUT 14 471–485	FFKVPETRGRTFEDI	42E
INS 48–61	FFYTPKTR—REAEDL	

Table 12. Na⁺/Glucose high affinity co transporter 1/Insulin similarities.

SLC5A1 597–610	KKKGIFRRAYDLFCGLEQ	39E
INS 87–105	QKRGIVEQCCTS I CSLYQL	
SC5A1 517–527	CPTIICGVHYL	36E
SLC5A1 175–191	LNL Y L A I FLLAI T A L Y	37E
INS 1–17	MALWMRLLPLLALLALW	
SLC5A1 276–318	WPGF IFGMS ILTL WYWCTDQVIVQRCLSAKNMSHVKGGCILCG	36E
INS 4–44	WMRLLPLLALLALWGPDPAAAFVNQHLCGSHL -- VEALYLVCG	
SLC5A1 184–218	LLAITALYTI T GGLAAVIYTD T L Q T V I M L -- V G S L I L	31E
SLC5A1 79–100	IGS GHFVGLAGTGAASGIA IGG	35E
INS 56–84	REAEDLQVGQVELGGGPGAGS L Q P L A L E G	
SLC5A1 603–618	R R A Y D L F C G - L E Q H G A P	33E
SLC5A1 166–184	A I F I N L A L - G L N L Y L A I F L L	34E
INS 20–42	GPDPAAAFVNQHLCGSHLVEALYL	
SLC5A1 464–472	GPP IAAVFL	33E

Table 13. Na⁺/Glucose low affinity co transporter 2/Insulin similarities.

SLC5A2 276–317	WPALLGLT IVSGWYWCSDQVIVQRCLAGKSLTHIKAGCILCG	45E
INS 4–43	WMRLPLLALLALWGPDPAAAFVNQHLCGSHL–VEALYLVCG	
SLC5A2 649–661	WARVVNLNALLMM	39E
SLC5A2 76–97	IG SGHFVGLAGTGAASGLAVAG	39E
INS 63–84	VGQVELGGGPGAGSLQPLALEG	
SLC5A2 558–617	RLVFSLRHSKEEREDLDADEQQGS SLPVQNGCPESAMEMNEPQAPAPSLFRQCLLWFCGM	38E
INS 46–102	RGFFYTPKTRREAEDLQVGVQVELGGGPGAGSLQPLALEGSLQKR --- GI VEQCCT S I C SL	
SLC5A2 621–631	GVGSPPLTQE	35E
INS 73–83	GAGSLQPLALE	
SLC5A2 67–89	VGASLFAS N IGSGHFVGLAGTGA	35E
INS 63–85	VGQVELGGGPGAGSLQPLALEGS	

A final test of the hypothesis is bioinformatic. All glucose transporters in all species should contain insulin-like sequences in their transport cores that are very highly conserved compared to other regions of the transporters. One of the interesting questions that arises evolutionarily is therefore whether insulin evolved prior to glucose transporters, parallel to them, or from some precursor of the glucose transporters. Currently available data do not permit this problem to be resolved. Determining the structural relationships of insulins and GLUT derived from a very wide range of organisms

may therefore shed important light on the molecular origins of glucose regulation and, in turn, new targets for the development of novel therapeutic regimens for diabetes.

Clinical Implications

The presence of insulin-like sequences in the GLUT has several practical implications for understanding the consequences of hyperglycemia in both type 1 and type 2 diabetes. Perhaps the most important follows from the possibility that the

Table 14. Na⁺/Glucose low affinity co transporter 3/Insulin similarities.

SLC5A4 276–317	WPG I IFGMP ITALWYWCTNQVIVQRCLCGKDMSHVKAACIMC	54E
INS 4–43	WMRLPLLALLALWGPDPAAAFVNQHLCGSHL -- VEALYLVCG	
SLC5A4 140–151	RLQVYLSILSLF	26E
SLC5A4 464–477	GPP I AAVFVLA I FC	42E
INS 17–31	GDPAAAFVNQHLC	
SLC5A4 133–143	KKRFGGERLQV	27E
INS 53–63	KTRREAEDLQV	
SLC5A4 599–621	LKKAYDLFCGLQKG ---- PKLTKEEEE	32E
INS 35–60	LVEALYLVCG- ERG FFYTPKTRREAED	

Table 15. Shared regions of homology between GLUT 1 and GLUT 4 and INSULIN, color coded as the key to Figure 1.

GLUT 1 56–93	I LPT–T LTTL W SLS – VA I FSVGG M I GSFSVGLFV NRFGRR	<u>WHITE</u>
INSULIN 7–46	LLPLLALLAL W GPDPAAAFV NQH LCGSHLVEALYLVCG ER	
GLUT 4 90–109	IPPG– TLTT W ALS – VA I FSVGG M I SSF L IG I ISQWLGRK	
GLUT 4 153–180	SGLVPMYVGE I APTHLRGALGTLNQLAI	<u>YELLOW</u>
INSULIN 56–82	REAEDLQVGQVELGGGPGA–GSLQPLAL	
GLUT 1 137–164	TGFVPMYVGEVSPTAFRGALGTLHQLG I	
GLUT 1 181–190	GN–KDL W PLLL	
GLUT 4 307–330	FYYSTS I FETAGVGQPAYAT IGAG	<u>PURPLE</u>
INSULIN 47–71	FFYTPKTRREAEDLQVGQVELGGG	
GLUT 1 291–315	FYYSTS I FEKAGVQQPVYATIGSG	
INSULIN 18–44	GPDPAAAFVNQHLGSHLVEALY LVCG	<u>RED</u>
GLUT 4 398–424	GPGPIPW FIVAELF SQGPRPAAMAVAG	
INSULIN 10–26	LLALLA LWGPDPAAAFV	
GLUT 1 390–406	I VAELFAQGPRPAA I AV	
GLUT 4 273–286	<u>LSLLQIL</u> GSR THR	<u>TURQUOISE</u>
INSULIN 79–92	<u>IQPLALEGSLQKR</u>	
GLUT 1 258–269	VT I LELFRS PAYR	
GLUT 4 124–147	SLMGLANAAASYEML ILGRF LIGA	<u>GREEN</u>
INS 15–38	ALWGPDPAAAFVNQHLGSHLVEA	
GLUT 1 112–135	VLMGFSKLGKSFEM LI LGRF I I GV	

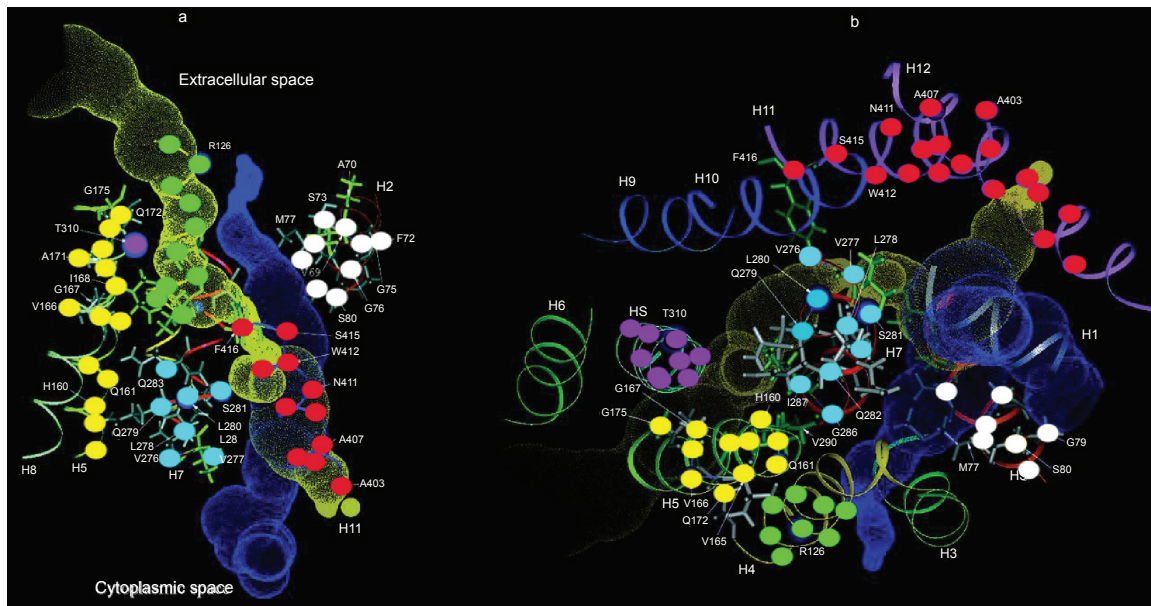


Figure 1. Regions of shared homology between GLUT 1, GLUT 4 and insulin (color coded in TABLE 15) mapped onto Zuniga et al.'s (2001) model of the GLUT transport core.

insulin-like regions of GLUT glycate rapidly under hyperglycemic conditions. (Parenthetically, one might expect fructose glycation of GLUT 2 and GLUT 5 since these transport fructose as well as glucose [Schalkwijk, Stehouwer and van Hinsbergh, 2004]). The glycation of binding sites for glucose transport would be predicted to interfere with glucose transport by blocking the transport cores, thus exacerbating hyperglycemia. The greater the hyperglycemia, the greater the glycation, and the greater the percentage of GLUTs that would malfunction. A positive feedback system would be set up that is self-destructive. The paradoxical result would be that the greater the hyperglycemia, the less glucose would be transported into cells. Cells, starved for a ready energy source, would have to metabolize protein, resulting in the ketosis associated with hyperglycemia. In children with diabetes, the result of such refractory glucose transport would be to inhibit critical developmental processes that are glucose-dependent, such as bone growth (Mobasher et al. 2002) and insulin-dependent neuron growth (Whitmer, 2007), resulting in long term deficits.

Two additional effects of hyperglycemic glycation might also result. First, if GLUT and sodium-dependent glucose transporters are critical sensors for blood glucose concentration as is currently thought (Freeman et al. 2006; Uldry and Thorens, 2004; Bogacka et al. 2004; Ferber et al. 1994;

Leloup et al. 1994), then glucose regulation might be impaired because glucose sensing would be impaired not only in somatic cells but in the brain. Various types of neurons act as glucose sensors in the brain, helping to regulate glucose homeostasis (Parton et al. 2007), so that blocking glucose transport in neurons would also exacerbate glucose dysregulation. Secondly, if the same rapid glycation reaction occurs on insulin-like regions of the insulin receptor (Root-Bernstein, 2002; Root-Bernstein, 2005), then insulin activation of its receptor may be impaired, resulting in the insulin-resistance that is associated with diabetes. The combination of these two malfunctions would be an inability for the glucose-regulatory system to sense or respond to its normal regulatory cues.

Understanding the evolutionary structure-function relationship of insulin and GLUT may also provide hints about a novel therapeutic approach to treating auto-glycation. Since GLUT transport dehydroascorbate (DHA), it is possible that serum ascorbate or DHA levels may be a critical factor in preventing (the assumed) GLUT auto-glycation. In essence, higher levels of DHA (which can be generated by higher levels of dietary or transfused ascorbate or DHA itself), will compete with glucose for the GLUT transport sites; DHA competition for glucose binding sites on insulin-like regions of GLUT may decrease auto-glycation of GLUT (as well as on insulin and the insulin

receptor), thereby helping to prevent some of the deleterious effects of hyperglycemia. This possibility is also testable by determining rates of auto-glycation of insulin and or insulin-like regions in GLUT in the presence and absence of various concentrations of DHA and ascorbic acid. The likelihood that this approach will be successful is indicated by a report by Abdel-Wahab et al. (2002) that ascorbic acid supplementation decreases insulin glycation in obese hyperglycemic mice. Clearly, these results need to be extended to determine rates of glycation of GLUT and insulin receptors in the presence of ascorbate and DHA.

Finally, the same factors just discussed in terms of diabetes may have relevance for understanding some aspects of normal aging as well. The rate of protein turn-over, particularly of GLUT, is well-documented to decrease with increasing age (Mooradian and Shah, 1997). This means that as a person ages, GLUT will be exposed for longer times to normal glucose serum concentrations. In addition, ascorbate (and thus DHA) deficiency increases with age (Junquiera et al. 2004) The result of prolonged exposure of GLUT to glucose and decreased antioxidant concentrations would be increased GLUT glycation, resulting in body-wide decreased sugar transport and decreased glucose sensitivity. These are early symptoms of metabolic syndrome, a pre-diabetic state also known to increase in prevalence with aging (Rodriguez et al. 2007; Yaffe, 2007). Indeed, the formation of advanced glycation endproducts is a major problem not only in diabetes, but in dementias as well (Vlassara and Palace, 2002; Baynes, 2001). Perhaps these syndromes, too, might be responsive to increased DHA intake.

Summary

In sum, I have proposed an evolutionary basis for GLUT structure based on a modular units originating in insulin-glucose binding. The model that results suggests various experimentally testable implications including a mechanism for GLUT transport of glucose via a series of insulin-like sequences that form the GLUT transport cores. The proposed structure also suggests the possibility that GLUT are auto-glycated under hyperglycemic conditions, resulting in defective glucose transport and regulation, and it proposes that dehydroascorbate (DHA) may prevent such auto-glycation by direct competition for the glucose binding sites.

Thus, a simple modular approach to GLUT evolution based on small molecule complementarity may provide basic insights into the structure and function of a major protein and its modifications in disease states.

Acknowledgements

I would like to thank my research assistant Sherika Gibson for help with the similarity searches. Two anonymous reviewers asked helpful questions that improved the manuscript. I also thank Maurine Bernstein for her generous donation of funds supporting this research in memory of Morton Ira Bernstein, who survived diabetes for many decades.

References

- Abdel-Wahab, Y.H.A., O'Harte, F.P.M., Barnett, C.R. and Flatt, P.R. 1997a. Characterization of insulin glycation in insulin-secreting cells maintained in tissue culture. *J. Endocrinol.*, 152:59–67.
- Abdel-Wahab, Y.H.A., O'Harte, F.P.M., Boyd, A.C., Barnett, C.R. and Flatt, P.R. 1997b. Glycation of insulin results in reduced biological activity in mice. *Acta Diabetol.*, 34:265–70.
- Abdel-Wahab, Y.H.A., O'Harte, F.P.M., Mooney, M.H., Barnett, C.R. and Flatt, P.R. 2002. Vitamin C supplementation decreases insulin glycation and improves glucose homeostasis in obese hyperglycemic (ob/ob) mice. *Metabolism*, 51:514–7.
- Anzenbacher, P. and Kalous, V. 1975. Binding of D-glucose to insulin. *Biochim. Biophys. Acta.*, 386:603–7.
- Baynes, J.W. 2001. The role of AGEs in aging: Causation or correlation. *Exp. Gerontol.*, 36:1527–37.
- Bogacka, I., Roane, D.S., Xi, X., Zhou, J., Li, B., Ryan, D.H. and Martin, R.J. 2004. Expression levels of genes likely involved in glucose sensing in obese Zucker rat brain. *Nutr. Neurosci.*, 7:67–74.
- Bornberg-Bauer, E., Beaussart, F., Kummerfeld, S., Teichmann, S. and Weiner, J. 3rd, 2005. The evolution of domain arrangements in protein and interaction networks. *Cell Mol. Life Sci.*, 62:435–45.
- Boyd, A.C., Abdel-Wahab, Y.H.A., McKillop, A., McNulty, H., Barnett, C.R., O'Harte, F.P.M. and Flatt, P.R. 2000. Impaired ability of glycosylated insulin to regulate plasma glucose and stimulate glucose transport and metabolism in mouse abdominal muscle. *Biochim. Biophys. Acta.*, 1523:128–34.
- Brownlee, M. 1995. Advanced protein glycosylation in diabetes and aging. *Ann. Rev. Med.*, 46:223–34.
- Dollhofer, R. and Wieland, O.H. 1979. Preparation and biological properties of glycosylated insulin. *FEBS Lett.*, 100:133–6.
- Farah, M.A., Bose, S., Lee, J.H., Jung, H.C. and Kim, Y. 2005. Analysis of glycosylated insulin by MALDI-TOF mass spectrometry. *Biochim. Biophys. Acta.*, 1725:269–82.
- Ferber, S., BeltrandelRio, H., Johnson, J.H., Noel, R.J., Cassidy, L.E., Clark, S., Becker, T.C., Hughes, S.D. and Newgard, C.B. 1994. GLUT2 gene transport into insulinoma cells confers both low and high affinity glucose-stimulated insulin release. Relationship to glucokinase activity. *J. Biol. Chem.*, 269:11523–9.
- Freeman, S.L., Bohan, D., Darcel, N. and Raybould, H.E. 2006. Luminal glucose sensing in the rat intestine has characteristics of a sodium-glucose cotransporter. *Am. J. Physiol. Gastrointest Liver Physiol.*, 291:G439–45.
- Hruz, P.W. and Mueckler, M.M. 2001. Structural analysis of the GLUT1 facilitative glucose transporter. *Mol. Mem. Biol.*, 18:183–93 (review).

- Huang, X. and Miller, W. 1991. A time-efficient, linear-space local similarity algorithm. *Adv. Appl. Math.*, 12:337–57.
- Hunding, A., Lancet, D., Kepes, F., Minsky, A., Norris, V., Raine, D., Sriram, K. and Root-Bernstein, R. 2006. Compositional complementarity and prebiotic ecosystems in the origin of life. *Bioessays*, 28 (4):399–412.
- Junqueira, V.B., Barros, S.B., Chan, S.S., Rodrigues, L., Giavarotti, L., Abud, R.L. and Deucher, G.P. 2004. Aging and oxidative stress. *Mol. Aspects Med.*, Feb–Apr; 25(1–2):5–16.
- Kalous, V. and Anzenbacher, P. 1979. On the mechanism of the insulin-glucose interactions. *Acta. Diabetol. Lat.*, 16:169–174.
- Leloup, C., Arluison, M., Lepetit, N., Carier, N., Marfaing-Jallat, P., Ferre, P. and Penicaud, L. 1994. Glucose transporter 2 (GLUT 2) expression in specific brain nuclei. *Brain Res.*, 638:221–226.
- Liu, J. and Rost, B. 2004. CHOP: Parsing proteins into structural domains. *Nucleic Acids res.*, 32:W569–W571.
- McKillop, A.M., Meade, A., Flatt, P.R. and O’Harte, F.P.M. 2003. Evaluation of the site(s) of glycation in human proinsulin by ion-trap LCQ-electrospray ionization mass spectrometry. *Regul. Pept.*, 113:1–8.
- Mobasheri, A., Vannucci, S.J., Bondy, C.A., Carter, S.D., Innes, J.F., Arteaga, M.F., Trujillo, E., Ferraz, I., Shakibaei, M. and MartinVasallo, P. 2002. Glucose transport and metabolism in chondrocytes: a key to understanding chondrogenesis, skeletal development and cartilage degradation in osteoarthritis. *Histol. Histopathol.*, 17(4):1239–67.
- Mooradian, A.D. and Shah, G.N. 1997. Age-related changes in glucose transporter-one mRNA structure and function. *Proc. Soc. Exp. Biol. Med.*, 216(3):380–5.
- O’Harte, F.P.M., Hojrup, P., Barnett, C.R. and Flatt, P.R. 1996. Identification of the site of glycation of human insulin. *Peptides*, 17:1323–30.
- O’Harte, F.P.M., Boyd, A.C., McKillop, A.M., Abdel-Whab, Y.H.A., McNulty, H., Barnett, C.R., Conlon, J.M., Hojrup, P. and Flatt, P.R. 2000. Structure, antihyperglycemic activity and cellular actions of a novel diglycated human insulin. *Peptides*, 21:1519–26.
- Olsowski, A., Monden, I., Kruase, G. and Keller, K. 2000. Cysteine scanning mutagenesis of helices 2 and 7 in GLUT1 identifies an exofacial cleft in both transmembrane segments. *Biochemistry*, 39:2469–75.
- Parton, L.E., Ye, C.P., Coppari, R., Enriori, P.J., Choi, B., Zhang, C-Y., Xu, C., Vianna, C.R., Balthasar, N., Le, C.E., Elmquist, J.K., Cowley, M.A. and Lowell, B.B. 2007. Glucose sensing by POMC neurons regulates glucose homeostasis and is impaired in obesity. *Nature*, 449:228–32.
- Patthy, L. 1999. Blackwell Science, Protein Evolution, Oxford.
- Rodriguez, A., Muller, D.C., Metter, E.J., Maggio, M., Harman, S.M., Blackman, M.R. and Andres, R. 2007. Aging, androgens, and the metabolic syndrome in a longitudinal study of aging. *J. Clin. Endocrinol. Metab.*, 92(9):3568–72.
- Root-Bernstein, R.S. and Dillon, P.F. 1997. Molecular complementarity, I: The molecular complementarity theory of the origin and evolution of life. *J. Theor. Biol.*, 188:447–479.
- Root-Bernstein, R.S. and Dobbstein, C. 2001. Insulin binds to glucagon forming a complex that is hyper-antigenic and inducing complementary antibodies having an idiotype-antiidiotype relationship. *Autoimmunity*, 33:153–69.
- Root-Bernstein, R.S., Busik, J. and Henry, D. 2002. Are diabetic neuropathy, retinopathy, and nephropathy caused by hyperglycemic exclusion of ascorbate uptake by glucose transporters. *J. Theor. Biol.*, 216:345–59.
- Root-Bernstein, R.S. 2002. Molecular complementarity III. Molecular complementarity as the basis for peptide hormone evolution. A bioinformatic case study of insulin, glucagon, and gastrin. *J. Theor. Biol.*, 218:71–84.
- Root-Bernstein, R.S. 2005. Peptide self-aggregation and peptide complementarity as bases for the evolution of peptide receptors: A review. *J. Molec. Recognit.*, 18:40–49.
- Rudensky, A.Y., Preston-Huriburt, P., Hoag, S-C., Balrow, A. and Janeway, C.A. 1991. Sequence analysis of peptides bound to MHC class II molecules. *Nature*, 353:622–7.
- Schalkwijk, C.G., Stehouwer, C.D.A. and van Hinsbergh, V.W.M. 2004. Fructose-mediated non-enzymatic glycation: sweet coupling or bad modification. *Diabetes Metab. Res. Rev.*, 20:369–82.
- Seatter, M.J., de la Rue, S.A., Porter, L.M. and Gould, G.W. 1998. QLS motif in transmembrane helix VII of the glucose transporter family interacts with the C-1 position of D-glucose and is involved in substrate selection at the exofacial binding site. *Biochemistry*, 37:1322–6.
- Uldry, M. and Thorens, B. 2004. The SLC2 family of facilitated hexose and polyol transporters. *Pflugers Arch.*, 447:480–9.
- Vlassara, H. and Palace, M.R. 2002. Diabetes and advanced glycation endproducts. *J. Intern. Med.*, 251:87–101.
- Whitmer, R.A. 2007. Type 2 diabetes and risk of cognitive impairment and dementia. *Current Neurol. Neurosci. Rep.*, 7:373–80.
- Yaffe, K. 2007. Metabolic syndrome and cognitive disorders: is the sum greater than its parts? *Alzheimer. Dis. Assoc. Disord.*, 21(2):167–71.
- Yu, G.B. and Caspar, D.L. 1998. Structure of cubic insulin crystals in glucose solutions. *Biophys. J.*, 74:616–22.
- Zoete, V., Meuwly, M. and Karplus, M. 2004. Investigation of glucose binding sites on insulin. *Proteins*, 55:568–81.
- Zuniga, F.A., Chi, G., Halber, J.F., Rubaskin A., Flynn, D.R., Isorovich, P., Fishbarg, J. 2001. *J. Biol. Chem.*, A three dimensional model of the human facilitative transporter Glut1, 276:4490–5.