ELSEVIER

Biophysical Chemistry 59 (1996) 95–105

Biophysical Chemistry

Electronic energy levels and hopping conductivity of an entire molecule of hagfish insulin

Yan Jiang, Yuan-Jie Ye *, Run-Sheng Chen

Laboratory of Protein Engineering, Institute of Biophysics, Chinese Academy of Science, 15 datun Road, Chaoyang District, Beijing 100101, China

Received 22 March 1995; revised 19 June 1995; accepted 10 August 1995

Abstract

The quantum-chemical calculation of an entire molecule of hagfish insulin was done by the ENFC method in which the matrix elements were calculated at the ab initio level using a minimal basis set with simulation of the aqueous solution environment. The ac conductivity for hagfish insulin was also calculated at the ab initio level by random walk theory. All the results were compared with those of pig insulin. It is shown that the reduction of HOMOs and LUMOs localized on the active sites of hagfish insulin agrees with the decrease in the biological reactivity of the insulin. The analysis of primary hopping events showed that a different sequence could influence the biological activity of insulin through the distribution of the hopping centers and the quantities of the hopping frequencies. The curve of the frequency versus ac conductivity of hagfish insulin shows that the different amino acid sequences of proteins influence the hopping conductivity. However, the electronic properties of native proteins are dominated by the three-dimensional conformations. Finally, the electronic mechanism of trans-membrane signal transforms by insulin and its receptor, which had been proposed by Ye and Ladik, was clearly described.

Keywords: Insulin; Pig; Hagfish; Proteins; Electronic energy level; Conductivity; Hopping; Signal transformation

1. Introduction

The reactivity of insulin in relation to its molecular structure has been investigated for decades [1–5]. As a hormone, insulin has various important biological functions and uses very complicated processes for exerting its biological activities. The relationship between the three-dimensional conformation and biological activities of the insulin molecule has been studied by standard approaches in terms of hydrogen bonds, hydrophobicity, and Van der Waals forces in previous works. In earlier studies, it was found that the part of the insulin molecule which forms a dimer in its crystal form, and the binding interactions of the insulin molecule to its receptor are important factors for explaining its activity [3]. Gammeltoft [3] considered the receptor binding area of insulin as a hydrophobic region surrounded by a hydrophillic one. This area was extended by Liang et al. [5] to include more residues being involved in the expression of insulin activity. As the experimental studies on in-

* Corresponding author.

SSDI 0301-4622(95)00116-6

sulin derivatives of varying biological activity and their structural comparisons were developed, more residues were found to be related to the explanation of its activity [6,7]. All these facts showed that the whole insulin molecule interacts with its receptor. That is, the binding to its receptor is a complicated process in which many residues of insulin are involved.

The structure–function relationship of insulin as a whole is important to explain its biological activity. Thus, the electronic structure of the entire molecule might also play a certain role in the expression of its biological reactivity. Therefore, it is necessary to investigate the electronic energy levels of the entire molecule. Recently, the electronic structure of proteins was investigated using periodic and aperiodic polypeptide models [8]. It was found that the activity of proteins and DNA may be explained partially by their electric conductivity. Ye [9–11] has proven that the extended negative factor counting (ENFC) theorem can solve the eigenvalue problem of tridiagonal block matrices with elements corresponding to cross-links, which may be derived for the quantum-chemical calculation on a native protein molecule.

The fact that proteins can become semiconductors rather than insulators was first reported by Szent–Györgyi [12] in 1941 and has now been explored both experimentally and theoretically. Ladik and coworkers [13–17] studied the conductivity of aperiodic models of proteins with the help of the negative factor counting method [18,19] and pointed out that the ac conductivity of proteins should be brought about by the hopping of the charge carriers instead of Bloch-type transport in the proteins.

As an application of the ENFC method, Ye and Ladik [20] have calculated the electronic energy levels of the entire molecule of pig insulin and studied its hopping conductivity [22,23]. It was shown that the frontier orbitals of the insulin were mainly localized on those residues which are involved in the expression of biological activity of insulin, and that the curve of the frequency versus ac conductivity of pig insulin lies in the range of some typical inorganic amorphous conductors and confirms that proteins, if doped, are amorphous conductors. The further calculation on the molecule II of pig insulin showed that the changes in the three-dimensional conformation of the insulin strongly influenced the ac conductivity of pig insulin [23].

The evident similarity in structure and preservation of sequence in the regions of the hormone considered to be involved in biological activity made us very interested in the electronic structure and the hopping conductivity of insulin. In this article, the electronic energy levels of the entire molecule and hopping conductivity for Atlantic hagfish (*Myxine glutinosa*) insulin were calculated by using the ab initio SCF LCAO method applying a minimal basis set for the ENFC method with the simulation of an aqueous solution surrounding the molecule. In the case of hagfish, the most primitive vertebrate and also the most primitive animal, its insulin sequence shows ca. 40% changes (19 changes in 51 residues), and the tertiary structure is similar to that of pig

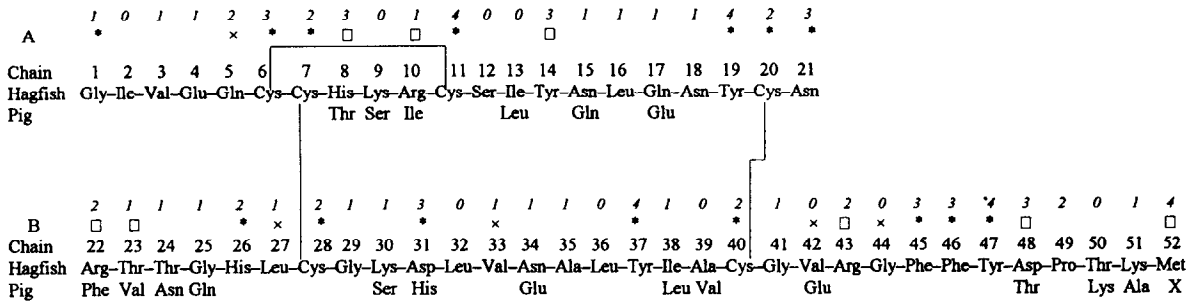

Fig. 1. The primary sequence of hagfish insulin, the relationship between its frontier orbitals and active sites and the distribution of the hopping centers which are taken into account in the calculation of its ac conductivity. (*) Active sites with HOMOs and LUMOs; (×) active sites without HOMOs and LUMOs; (□) uncertain sites with HOMOs or LUMOs. The *italic numbers* given above the different residues indicate the number of hopping centers of the different orbitals localized mainly on the residue. The residues displayed under the sequence are those of pig insulin, which are different from those of hagfish insulin.

insulin [24]. There are, however, substantial differences in their biological activity and binding affinity. The values of its biological activity and binding affinity relative to pig insulin determined in an isolated rat fat cell assay are 7% and 25%, respectively. Most natural and synthetic insulin analogues exhibit a binding affinity relative to pig insulin that is identical to the relative biological activity. Hagfish insulin, however, does not confirm this rule. Its relative biological activity is more than 3 times lower than its binding affinity in fat cell. Its decreased efficacy may indicate that binding and activation functions of the insulin molecule are at least partially isolated [3]. Therefore, the hagfish insulin is a good example to investigate the influence of amino acid sequence on the electronic structures and ac conductivity of insulin.

In this paper we shall report the results of the electronic structures and the hopping conductivity of an entire molecule of hagfish insulin. In the next section we briefly introduce the approximations and methods used in the calculations. In the third section we report and discuss the calculated results. The conclusions are presented in the last section and some deductions are discussed in the same section.

2. Approximations and methods

Fig. 1 shows the primary sequence of hagfish insulin. The three-dimensional structural data [24] of the insulin were used in the calculation. The molecular crystal has one molecule per asymmetric unit but is organized as a symmetric dimer lying on a 2-fold crystal axis. All of the coordinates of the hydrogen atoms which were not listed in the original data set were obtained theoretically. There are 52 residues which form two peptide chains, two disulphide bridges between the two chains and another disulphide bridge in chain A, and 811 atoms, including hydrogen atoms, in hagfish insulin. The total number of the basis functions is 2487 when a minimal basis set is used.

It is impossible to work out the electronic energy levels of an entire protein molecule by full SCF method in an ab initio scheme because it needs quite a lot of computer memory capacity (more than 2000 Gbytes) to store two-electronic integrals of the whole system. Therefore, one has to do some approximations in such calculations. Ye and Ladik have suggested the overlap cluster approximation [11,20] to solve this problem. In this approximation the Fock matrix of a whole molecule can be constructed with the help of the Fock matrices of clusters only corresponding to the interactions of first neighbours that are close to each other in the three-dimensional conformation. In this paper, these clusters, in order to simplify the calculation, are constructed by two amino acid residues that are linked by a chemical bond, that is, peptide bond or disulphide bridge. At the end of each dimer pseudo-atoms are added to simulate the chemical environment of the other nearest neighbouring units of the dimer. The Fock matrix constructed in this way is very close to that of the full SCF method [25]. Therefore, the electronic energy levels obtained are assumed equal to those for the entire molecule. The wave functions which belong to these energy levels are also calculated for the entire molecule, that is, all 2487 basis functions are taken into account in the calculations of the wave functions. In this way, a wave function would be distributed all over the molecule if it were fully delocalized (for details, see Ref. [20]).

One has to simulate an aqueous solution environment surrounding the protein molecule in the quantum chemical calculation. First, under normal conditions, protein crystals are prepared in an aqueous solution and there are some water molecules in the protein crystals. Second, it has been known that insulin molecules express their biological activities in the aqueous solution. To simulate the solution environment [20], point charges were put around the residues that have electronic charges (see Fig. 3 of Ref. [20]). For comparison with the results of pig insulin, we take the same scheme as that for pig insulin. The details of the simulation have been reported in Ref. [20].

Under the conditions of this simulation, i.e. at pH 7.0, each of the Glu and Asp amino acids has a negative charge at its side chain because they are not protonated. Each side chain of Lys and Arg is protonated and has a positive charge. His is neutral and has a δ proton but not an ε proton [21]. The N-terminal has a positive charge, and the C-terminal has a negative charge for both chains.

The results of hopping conductivity of hagfish

insulin were obtained by using the energy levels and molecular orbitals which had been calculated. The physical model and approximations taken in this paper were the same as those of the previous works [22,23,26] in which the results of hopping conductivity of pig insulin and lysozyme have been reported. For the sake of the readers we briefly introduce them in this paper. We define the hopping centers and "main residue" of a molecular orbital, and assume that the hopping of charge carriers occur among these centers of main residues. In the calculation only the hoppings among the centers of the same residue and between nearest and next neighbouring residues are taken into account (in this definition hopping through the disulphide bridges should be included, see Ref. [22]). The hopping frequencies between the hopping centers are calculated by the same formulae as in Ref. [22]. The protein molecule was considered as an amorphous system and the random walking theory of hopping conductivity [27–30] is applied to calculate its conductivity (for detail, see Refs. [22,26]). In this article hagfish insulin is also considered as a quasi-one-dimensional system although its three-dimensional conformation is used in the calculation of hopping conductivity. To compare the results with those of pig insulin, we take the 50 highest filled and 50 lowest unfilled orbitals in the calculation of the ac conductivity of hagfish insulin too. They were only used in the construction of the hopping matrix **H** [22]. Thus, the number of charge carriers is 100 and the molecular volume is 20960.0 Å^3 [24] for hagfish insulin. The frequency range is taken from 1 to 10^{20}.

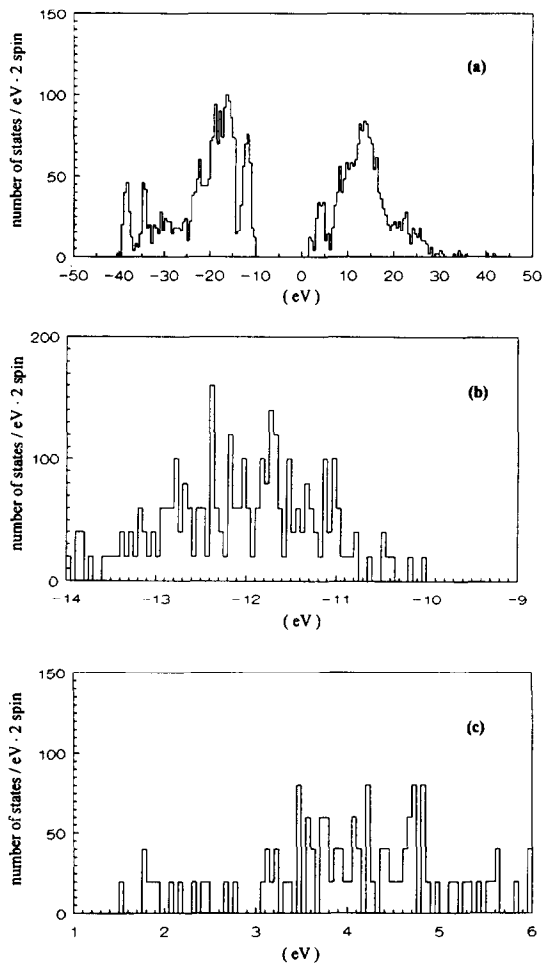

Fig. 2. The electronic DOS of hagfish insulin: (a) both valence and conduction band regions using a grid of 0.5 eV; (b) the valence band region using a grid of 0.05 eV; (c) the conduction band region using a grid of 0.05 eV. (The grids are not the scales shown on the abscissa, but the step lengths in the DOS histograms.)

3. Calculated results

3.1. Density of states

The electronic density of states (DOS) of hagfish insulin is shown in Fig. 2a–c, in which (a) is for the valence and conduction band regions applying a grid of 0.5 eV, while (b) and (c) are for parts of the valence and conduction band regions using a grid of 0.05 eV, respectively.

From Fig. 2 it can be seen that the DOS (in number of states/eV · 2 spin) of hagfish insulin is in good agreement with that of pig insulin [20], the band structures of both molecules being also similar to each other, and the highest filled peak in the DOS of the valence band regions of both molecules corresponding to the delocalized orbitals on the peptide backbone. The energy gap of hagfish insulin (11.55 eV) is slightly smaller than that of pig insulin (11.59 eV). These results indicate that the difference between the sequences of hagfish insulin and pig insulin does not seriously influence their electronic densities of states.

The calculated energy gaps (forbidden gaps) of

the two different kinds of insulin are so large that in their undoped states they should be insulators. In minimal basis set Hartree–Fock calculations on protein molecules one obtains an unrealistically large gap. If one would apply a far better basis set and would correct the band structure for correlation [31], the gaps can be estimated to be between 6 and 7 eV [32]. The experimental energy gap of lysozyme is about 3–4 eV [33]. These values are still far too large to assume intrinsic (thermal) conductivity in proteins. Thermal energy cannot pump any electron from the filled valence band's region to the empty conductive band's region. The conditions for proteins to become amorphous conductors are that they should lose or obtain electrons to create effective charge carriers in the valence and conductive band's regions, respectively. This can happen through biochemical reactions or photoexcitation. The conductivity of protein molecules is caused by hopping of these effective charges between different energy lev-

Table 1
Highest occupied molecular orbitals (HOMOs) of hagfish insulin

		Energy level (eV)	Position of the wavefunction (residue)	
1580	(1)	−10.04691	Asn^{21}	*
1579	(2)	−10.15027	Cys^{6}–Cys^{11}	*
1578	(3)	−10.35451	Asp^{48}	
1577	(4)	−10.43877	Met^{52}	
1576	(5)	−10.48934	Tyr^{37}	*
1575	(6)	−10.49857	*Tyr^{47}*–Asp^{48}	*
1574	(7)	−10.63786	Tyr^{19}	*
1573	(8)	−10.76130	Tyr^{47}–*Asp^{48}*–Pro^{49}	
1572	(9)	−10.79724	Lys^{30}–*Asp^{31}*–Leu^{32}	p.b.
1571	(10)	−10.82887	Cys^{20}–*Asn^{21}*	*
1570	(11)	−10.87337	Asp^{31}	*
1569	(12)	−10.93936	Cys^{20}–Cys^{40}	*
1568	(13)	−10.95022	Met^{52}	
1567	(14)	−10.97021	Ile^{2}–Val^{3}–Glu^{4}	p.b.
1566	(15)	−10.98334	His^{8}–Cys^{7}–Cys^{28}	*
1565	(16)	−11.00731	Ser^{12}–Iso^{13}–Tyr^{14}–Asn^{15}–Leu^{16}	p.b.
1564	(17)	−11.03285	Lys^{30}–Asp–^{31}–Leu^{32}	p.b.
1563	(18)	−11.04195	Tyr^{14}–Asn^{15}–...–Asn^{18}–Tyr^{19}	p.b.
1562	(19)	−11.04606	Iso^{13}–Tyr^{14}–Asn^{15}–...–Asn^{18}–Tyr^{19}	p.b.
1561	(20)	−11.04779	Asp^{48}–Pro^{49}–Thr^{50}	p.b.
1560	(21)	−11.05121	Leu^{36}–Tyr^{37}–Iso^{38}–Ala^{39}	p.b.
1559	(22)	−11.05999	His^{8}–Cys^{7}–Cys^{28}	*
1558	(23)	−11.10392	Glu^{4}–Gln^{5}–...–Cys^{7}–His^{8}–Cys^{28}	p.b.
1557	(24)	−11.11296	Glu^{4}–Gln^{5}–...–Cys^{7}–His^{8}–Lys^{9}	p.b.
1556	(25)	−11.11982	Pro^{49}–...–Lys^{51}–*Met^{52}*	
1555	(26)	−11.14453	Cys^{7}–His^{8}–Cys^{28}	
1554	(27)	−11.14690	Arg^{43}–Gly^{44}–Phe^{45}	p.b.
1553	(28)	−11.18607	Iso^{13}–Tyr^{14}–Asn^{15}–Leu^{16}–Gln^{17}	p.b.
1552	(29)	−11.20329	Asn^{34}–Ala^{35}–Leu^{36}	p.b.
1551	(30)	−11.22039	*Asn^{21}*–Cys^{20}–Cys^{40}	

The definition of a component at the nth residue is $a(n) = \sum_{j-1}^{m_n} C_j^2(n) / \sum_{n-1}^{N} \sum_{j-1}^{m_n} C_j^2(n)$.

Those residues with $a(n) < 0.05$ are neglected.

If the wavefunction belonging to a certain level is localized on several units, its main component belongs to the residue printed in *italics*. The definition of a main component is $a(n) \geq 0.60$.

Wave functions corresponding to an active site are indicated by *.

Wavefunctions mainly distributed at the peptide bonds between residues are indicated by p.b.

els in both the valence band and the conductive band, respectively (see Ref. [22] for details).

3.2. Frontier orbitals and active sites

The frontier orbitals (HOMOs and LUMOs) are important in chemical reactions. In a previous paper [20] the frontier orbitals of pig insulin have been reported. It has been shown that the distributions of the frontier orbitals are in good agreement with the active sites of pig insulin. The energy levels and positions of these frontier orbitals may be one of the factors to control the activity of insulin. The comparison of the frontier orbitals of hagfish insulin with those of pig insulin will be shown below.

Tables 1 and 2 list a part of the frontier orbitals of the hagfish insulin. Comparing Tables 1 and 2 in this paper with those of pig insulin [20] we note that there is a distinct difference in the distributions of the frontier orbitals between hagfish and pig insulins.

It can be observed that from the first 18 HOMOs, there are only 9 HOMOs corresponding to active

Table 2
Lowest unoccupied molecular orbitals (LUMOs) of hagfish insulin

		Energy level (eV)	Position of the wavefunction (residues)	
1581	(1′)	1.50331	*Tyr*^{14}	
1582	(2′)	1.75086	*Phe*^{45}	*
1583	(3′)	1.75668	Tyr^{19}	*
1584	(4′)	1.82201	Tyr^{47}	*
1585	(5′)	1.89188	Tyr^{37}	*
1586	(6′)	1.91350	Phe^{46}	*
1587	(7′)	2.06587	Phe^{46}	*
1588	(8′)	2.15037	Tyr^{14}	
1589	(9′)	2.34652	Tyr^{37}	*
1590	(10′)	2.41189	Phe^{45}	*
1591	(11′)	2.49430	Tyr^{19}	
1592	(12′)	2.65117	Tyr^{47}	*
1593	(13′)	2.76689	Cys^{6}–...–Cys^{11}	*
1594	(14′)	3.08622	Arg^{22}–*Thr*^{23}–Thr^{24}	
1595	(15′)	3.12129	Arg^{10}	
1596	(16′)	3.14740	Leu^{16}–Gln^{17}–Asn^{18}–Tyr^{19}–Cys^{20}	p.b.
1597	(17′)	3.18731	Cys^{7}–...–His^{26}–Leu^{27}–Cys^{28}	*
1598	(18′)	3.21658	*Gly*^{1}–Ile^{2}–Val^{3}	*
1599	(19′)	3.24826	His^{26}–...–Cys^{7}–Cys^{28}	*
1600	(20′)	3.30753	Arg^{43}	
1601	(21′)	3.35332	Phe^{46}–Tyr^{47}–Asp^{48}	p.b.
1602	(22′)	3.46417	Lys^{51}–*Met*^{52}	
1603	(23′)	3.49287	Cys^{6}–...–Leu^{16}–Gln^{17}–Asn^{18}–Tyr^{19}	p.b.
1604	(24′)	3.49297	Cys^{6}–Cys^{7}–His^{8}–...–Leu^{16}	p.b.
1605	(25′)	3.49594	Thr^{24}–Gly^{25}–His^{26}	p.b.
1606	(26′)	3.57524	Asn^{34}–Ala^{35}–Leu^{36}	p.b.
1607	(27′)	3.58671	Cys^{28}–Gly^{29}–Lys^{30}–Asp^{31}–...–Asn^{34}	p.b.
1608	(28′)	3.59089	Arg^{22}	
1609	(29′)	3.60159	Cys^{20}–Cys^{40}	
1610	(30′)	3.60830	Pro^{49}–Thr^{50}	p.b.

The definition of a component at the nth residue is $\alpha(n) = \sum_{j=1}^{m_n} C_j^2(n) / \sum_{n=1}^{N} \sum_{j=1}^{m_n} C_j^2(n)$.

Those residues that have $\alpha(n) < 0.05$ are neglected.

If the wavefunction belonging to a certain level is localized on several units, its main component belongs to the residue printed in *italics*. The definition of a main component is $\alpha(n) \geq 0.60$.

Wavefunctions corresponds to an active site are indicated by *.

Wavefunctions is mainly distributed at the peptide bonds between residues are indicated by p.b.

sites while among the first 22 LUMOs there are only 13 such orbitals, which is less than in pig insulin (13 out of the 18 HOMOs and 16 out of the 22 LUMOs, respectively). Therefore, one can say that the decrease in the sites of HOMOs and LUMOs of hagfish insulin relative to those of pig insulin causes the reduced activity of hagfish insulin.

It can be found that at the residue Phe^{46} (Phe^{B25}) in hagfish insulin, one of the most important residues for expression of biological activity of insulin, there is no high-lying occupied electronic energy level. Therefore, the chemical activity of the residue Phe^{46} (Phe^{B25}) of hagfish insulin is much lower than that of pig insulin. The energy level of the HOMO localized on the residue Tyr^{47} (Tyr^{B26}) of hagfish insulin is 0.315 eV (7.253 kcal/mol) higher than that of pig insulin so that the residue Tyr^{47} (Tyr^{B26}) should more active in hagfish insulin than that in pig insulin.

There is no MO with a filled high-lying energy level localized on the residue Val^{42} (Val^{B21}) of hagfish insulin's B chain. This suggests that this residue is chemically inert. However, the residue Glu^{42} (Glu^{B21}) of pig insulin, localized on the same site as in hagfish insulin, has a stronger chemical activity because it has a pair of electrons whose energy level lies at the first HOMO. The energy level of the HOMO of the residue Asp^{31} (Asp^{B10}) of hagfish insulin is 0.189 eV (4.348 kcal/mol) higher than that of His^{31} (His^{B10}) of pig insulin, and so is the energy level of the HOMO of the A chain residue His^{8} (His^{A8}) of hagfish insulin 0.283 eV (6.522 kcal/mol) higher than that of Thr^{A8} in pig insulin, which is a residue nearly like that found in turkey insulin, whose activity is higher than that of pig insulin [3]. This could mean that the residues Asp^{31} (Asp^{B10}) and His^{8} (His^{A8}) of hagfish insulin have stronger chemical activities than those of pig insulin.

Differences in the energy levels of three disulphide bridges also clearly exist. For the molecular orbital localized at the disulphide bridge Cys^{6}–Cys^{11} (Cys^{A6}–Cys^{A11}), there is evidence for an elevation of the energy level of 0.613 eV (14.138 kcal/mol) from the eighth HOMO of pig insulin to the second HOMO of hagfish insulin. The energy level of the disulphide bridge Cys^{7}–Cys^{28} (Cys^{A7}–Cys^{B7}), which lies in the fifteenth HOMO of hagfish insulin, however, is 0.155 eV (3.575 kcal/mol) lower than that of pig insulin. There are also evident energy level changes of the LUMOs localized at the disulphide bridges Cys^{20}–Cys^{40} (Cys^{A20}–Cys^{B19}) and Cys^{7}–Cys^{28} (Cys^{A7}–Cys^{B7}). The energy levels of the orbitals at these two disulphide bridges of hagfish insulin are 0.918 eV (21.166 kcal/mol) and 0.474 eV (10.924 kcal/mol), respectively, higher than those of pig insulin.

All these differences mentioned above imply that the relationship between the electronic structures and biological activity of insulin is subtle. Possibly hagfish insulin's reduced activity relative to pig insulin's activity is only partially due to the decrease in its sites of HOMOs and LUMOs corresponding to the active, sensitive sites. Even though some residues in hagfish insulin have strong chemical activities, others which have lower chemical activities relative to pig insulin may counteract their effects on insulin's activity. From the results shown in this paper one can hypothesize that the electronic energy levels of the frontier orbitals localized at the active, sensitive sites of insulin might also be responsible for the difference in the biological activity between hagfish and pig insulin. The lack of electrons that have high-lying energy levels at residues Phe^{46} (Phe^{B25}) and Val^{42} (Val^{B21}) of hagfish insulin might reduce its biological activity, although the electrons that have higher energy levels at the residue His^{8} (His^{A8}) could increase its biological activity. However, the rising of the energy level of the filled frontier orbital, which is localized at residue Asp^{31} (Asp^{B10}) of hagfish insulin, seems not enough to explain its effect on the high potency of Asp^{B10}–human insulin derivative [6]. Its effect might result from the negative charge of the side chain instead of its electronic energy level of the frontier orbital.

3.3. Primary hopping frequencies

Ye and Ladik [23] have proposed for the electronic mechanism of trans-membrane signal transformation that the insulin would change the ac conductivity of itself and of its receptor in the high-frequency range to transform the signal across a membrane when bound to its receptor. The primary hopping frequencies are the basic quantities for the calculations of the ac conductivity [22]. Therefore,

comparing the differences of the hopping frequencies between the two kinds of insulins will help us to analyze the electronic transport mechanism in detail.

The number of hopping centers taken into account in the calculation of ac conductivity are shown in Fig. 1. Combining this with the results of pig insulin (both Figs. 1 in Ref. [22] and in Ref. [23]), we can find that there are three regions that have more centers than the other parts. Two of them are around the disulphide bridge Cys^{20}–Cys^{40} (Cys^{A20}–Cys^{B19}) at the C-terminal of the A chain and around the Phe^{46} (Phe^{B25}), respectively, and they are close to each other in the three-dimensional structure and are included in Liang's proposed binding interaction surface of the insulin molecule with its receptor [5]. Another is around the two disulphide bridges Cys^{6}–Cys^{11} (Cys^{A6}–Cys^{A11}) and Cys^{7}–Cys^{28} (Cys^{A7}–Cys^{B7}). This fact indicates that the electronic transport mainly happens in these three regions. That is, the three regions should be the main paths in the electronic transport in an insulin molecule, and probably, in the insulin–receptor complex.

The largest hopping frequencies of different types (hopping between different orbitals localized on the same residues, between nearest and second-nearest neighbours, hopping through disulphide bridges) are presented in Table 3. Comparing them to those of pig insulin given in Table 2 of Ref. [22], we find that, in the results of the two different kinds of insulin, there exist quantitative differences of about one or two orders of magnitude, and the hopping events also happen at different sites in the two different insulin molecules.

From Table 3 we can see that the hopping frequencies between different orbitals localized on the same residues of hagfish insulin are quite different

Table 3
Some primary hopping frequencies of hagfish insulin

i	j	n	n'	ΔE_{ij} (eV)	$h_{X(n,i)\to X(n',j)}$	$h_{X(n',j)\to X(n,i)}$
40	33	26	26	0.10266	5.78×10^{8}	2.71×10^{10}
24	15	8	8	0.12962	6.45×10^{7}	8.29×10^{9}
11	9	31	31	0.07613	5.32×10^{7}	9.23×10^{8}
17	11	31	31	0.15948	8.92×10^{6}	3.51×10^{9}
25	13	52	52	0.16960	1.90×10^{6}	1.09×10^{9}
34 *	37 *	28	27	0.07081	3.09×10^{7}	4.39×10^{8}
27 *	34 *	29	28	0.17298	1.05×10^{6}	6.83×10^{8}
24	22	8	7	0.05297	4.29×10^{5}	3.12×10^{6}
31 *	41 *	5	4	0.21027	1.19×10^{5}	3.14×10^{8}
18	12	19	20	0.10260	7.33×10^{4}	3.42×10^{6}
32 *	35 *	41	43	0.04906	1.25×10^{5}	7.87×10^{5}
33	26	26	28	0.15533	2.84×10^{4}	9.57×10^{6}
40 *	45 *	17	15	0.14786	1.33×10^{4}	3.40×10^{6}
2 *	4 *	45	47	0.07114	7.82×10^{2}	1.12×10^{4}
31	28	14	16	0.09377	2.19×10^{2}	7.35×10^{3}
33 *	54 *	8	28	0.01097	1.12×10^{3}	1.68×10^{3}
18	12	19	40	0.10260	8.17×10^{2}	3.82×10^{3}
12	10	40	21	0.11049	5.84×10^{1}	3.67×10^{3}
29 *	32 *	20	41	0.13647	5.76	9.57×10^{2}
26	24	28	8	0.03157	4.76	1.55×10^{1}
26	22	28	7	0.08454	2.98×10^{6}	7.08×10^{7}
39	26	7	28	0.23251	2.33×10^{3}	1.41×10^{7}
24 *	43 *	6	11	0.51736	6.26×10^{-3}	1.64×10^{6}
13 *	24 *	11	6	0.72608	4.22×10^{-5}	2.75×10^{7}
13 *	43 *	6	11	1.24343	4.39×10^{-12}	7.50×10^{8}

i and j indicate the numbering of molecular orbitals; n and n' indicate the numbering of residues; $\Delta E_{ij} = E_j - E_i$ indicates the difference between the ith and jth energy levels; $h_{X(n,i)\to X(n',j)}$ indicates the hopping frequency of a charge carrier from the center at the nth residue of the ith orbital to the center at the n'th residue of the jth orbital; $h_{X(n',j)\to X(n,i)}$ indicates the hopping frequency of that charge carrier from the center at the n'th residue of the jth orbital to the center at the nth residue of the ith orbital (see formula 3 in Ref. [20]). Asterisks indicate unfilled energy levels whose numbering begins at the lowest one.

from those of pig insulin [22,23]. The hopping frequencies between nearest and next neighbouring residues in the same chains of hagfish insulin are also quite different from those of pig insulin. These facts indicate that the hopping mechanism is in detail different for the two kinds of insulin. This might be one of the causes of the reduction of biological activity of hagfish insulin.

In the hopping frequencies listed in Table 3 we do not find any one that can be ascribed to the residue Phe^{46} (Phe^{B25}) and find only one that belongs to Phe^{45} (Phe^{B24}) and Tyr^{47} (Tyr^{B26}) and which is much smaller than those appearing at the same sites of pig insulin [22]. This suggests that the contribution of hopping to electronic transport in this region of hagfish insulin is reduced relative to that of pig insulin because the probability of electron hopping is reduced. The evident change in the differences of energy levels of the disulphide bridges which have the largest hopping frequencies between hagfish and pig insulin suggests that the capability of electron transport in this region in hagfish insulin is also weaker than in pig insulin because the differences in energy levels of these disulphide bridges belonging to hagfish insulin are greater than those of pig insulin. According to the electronic mechanism of signal transformation by insulin and its receptor [23], we suggest that the low ability of electronic transformation in these regions might cause low activity of hagfish insulin.

From the calculated results of the primary hopping frequencies, we found that both insulins have a common feature that the residue sites which belong to a larger number of centers are in good agreement with the active sites that have been experimentally determined. Therefore, the binding affinity and the biological activity of insulins are closely related to each other because the binding regions contain the main electronic transport paths. The partial isolation of the binding affinity and the activation functions of hagfish insulin is probably caused by the lower ability of electronic transport in this insulin.

3.4. Hopping conductivity of hagfish insulin

The atoms which are involved in the biochemical reactions change their positions in a very short time period (picosecond). This time period corresponds to the high-frequency range of conductivity ($\omega > 10^{10}$ s^{-1}). The change of the positions of the atoms is, however, irreversible. Therefore, the treatment of a high-frequency electric current can be used as an approximation to describe such processes (for details, see Ref. [23]). Our interests in the following analysis will be focused on the high-frequency range of ac conductivity.

The ac conductivity calculated for hagfish insulin is presented in Fig. 3. In order to compare this with that of pig insulin, the calculated ac conductivity of pig insulin is also shown in the same figure. One can find that the hopping conductivity of hagfish insulin is qualitatively in agreement with that of pig insulin. However, there are, to a certain extent, some quantitative differences between these curves.

In the figure we can see that the real parts of the ac conductivity are more sensitive to the different kinds of insulin and about one order of magnitude different from each other. The combination of the real and imaginary parts makes the curves of the absolute value of the ac conductivity of the two insulins smoother. Except for a curve bend at $\omega \approx 10^4$ for hagfish insulin and at $\omega \approx 10^6$ for pig insulin, they all reach a constant value after $\omega \approx 10^{10}$. The constant value of hagfish insulin is about one order of magnitude smaller than that of molecule I of pig insulin but larger than that of molecule II of pig insulin.

The calculated results of the ac conductivity of hagfish insulin show that the 40% differences between the sequences of the two different kinds of insulin seem not to seriously influence their ac conductivity. The difference is smaller than that caused by the three-dimensional changes of pig insulin itself (see Fig. 3 and Ref. [23]). Therefore, the 40% difference in amino acid sequence between the two insulin molecules has less influence on the ac conductivity than the differences of the different three-dimensional conformations of pig insulin itself. Obviously, both insulins express their biological activity in the same way. The difference in their activity is caused by the different hopping mechanism in detail. Finally, we can conclude that the ac conductivity of proteins is dominated by three-dimensional conformations of protein molecules instead of amino acid sequences.

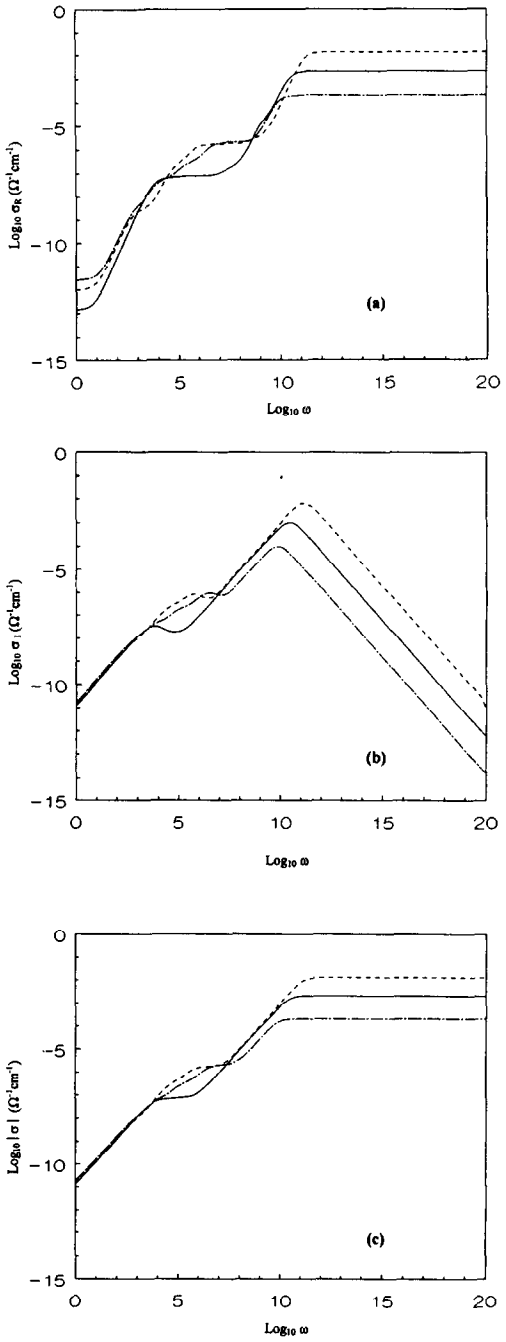

Fig. 3. The ac conductivities of hagfish and pig insulins. The solid line (—) represents that of hagfish insulin. The dashed line (– – –) and the dashed and dotted line (– · –) represent those of molecule I and II of pig insulin, respectively. (a) The real parts of the conductivities; (b) their imaginary part; (c) their absolute value.

4. Discussions and conclusions

In this paper we have investigated the electronic structures and hopping conductivity of hagfish insulin theoretically and compared the results to those of pig insulin. Summing up the all results we can conclude that the different amino acid sequences of the proteins influence their electronic structures and electronic transport. However, the electronic structures and electronic transport of a native protein are dominated by its three-dimensional conformations.

Up to now, the biological activities of insulins have been researched by standard approaches in terms of hydrogen bonds, hydrophobicity, and Van der Waals forces. The question remains of what kind of role the electronic structures, which are also dominated by three-dimensional conformations, play in the expression of the biological activity of insulins? A physical model has been proposed by Ye and Ladik [23] to answer this question. It was hypothesized that the insulin would change the ac conductivity of itself and of its receptor in the frequency range corresponding to the time period of chemical reactions when it is bound to its receptor. That is, the insulin molecule acts like an electronic switch. The signal transformation across the cell membrane by the receptor of insulin may be caused by electronic transport which was controlled by insulin. The results of this paper support this hypothesis that the electronic transport could be involved in the expression of biological activity of insulin because there are many detailed differences in the electronic structures and primary hopping events between the two different kinds of insulin. Moreover, the electronic transport might have more effect on the biological activity than on the binding affinity, because the latter may be dominated by non-bonding interactions except a few covalent interactions [34]. Therefore, the differences in electronic structures and primary hopping events between the two insulins might be one of the reasons that hagfish insulin has a different binding affinity and biological activity.

Combining the results of this paper with all the results of previous papers [10,20,22,23,26], we could describe the physical model of the electronic mechanism of signal transformation across a membrane by insulin and its receptor clearly. The insulin binds to the external membrane side of its receptor and

changes the three-dimensional conformation of itself and of its receptor. In this step the three-dimensional conformation of the external membrane side of the receptor will be changed and those of the other parts may also be influenced. The insulin–receptor complex has an ac conductivity which is largely different from their separate forms in the high-frequency range that corresponds to the time period of an elementary step of chemical reactions. It is probably one or two orders of magnitude higher according to the results of the calculations on both molecules of pig insulin. Therefore, it can be concluded that the complex forms an electronic channel across the membrane. The electrons could be transported through the channel. The effective charge carriers of the channel were created by the biochemical reactions occurring in the complex. At the same time only those electrons which are involved in the reactions could be transported because in a time period that is longer than an elementary step of chemical reactions the conductivity of the complex is still unchanged and much lower. The channel would then be closed and the biochemical reactions would stop when the insulin molecule is isolated from its receptor.

The three-dimensional conformations of the receptor and its complex with insulin are not yet known. We cannot verify this physical model by direct theoretical calculations. However, one could design an experiment to verify it. An alternating current across a membrane with high frequency (about 10^{11} Hz) could be measured after insulin binds to its receptor on the membrane. The effective charge carrier could be created by photoexcitation or biochemical reactions. This alternating current would become much weaker after the insulin is isolated from its receptor. Obviously, one should perform such experiment very carefully. The samples should be pure enough to avoid interferences with the measurements.

Acknowledgements

This project is supported by the National Scientific Foundation of China, and the "Fund for Returned Scientists" from the Chinese National Education Committee and Chinese Academy of Science, respectively.

References

[1] T. Blundell, G. Dodson and D. Hodgkin, Adv. Prot. Chem., 26 (1972) 279.
[2] D. Brandenberg and Z. Wollmer, Insulin, Chemistry, Structure and Function of Insulin and Related Hormones, De Gruyter, Berlin, 1980.
[3] S. Gammeltoft, Physiol. Rev., 64 (1984) 1321.
[4] E.N. Baker, T.L. Blundell, J.F. Cutfield, S.M. Cutfield, E.J. Dodson, G.G. Dodson, D.M.C. Hodgkin, R.E. Hubbard, N.W. Isaacs, C.D. Reynolds, K. Sakabe, N. Sakabe and N.M. Vijayan, Phil. Trans. R. Soc., 319 (1988) 369.
[5] D.-C. Liang, W. Chang and Z. Wan, Biophys. Chem., 50 (1994) 63.
[6] G.P. Schwartz, G.T. Burke and P.G. Katsoyannis, Proc. Natl. Acad. Sci. USA, 84 (1987) 6408; 86 (1989) 458.
[7] S.H. Nakagawa and H.S. Tager, J. Biol. Chem., 266 (1991) 11502.
[8] J. Ladik, Int. J. Quantum Chem., 23 (1983) 1073.
[9] Y.-J. Ye, Acta Phys. Chim. Sinica, 7 (1991) 257.
[10] Y.-J. Ye, Chem. J. Chin. Univ., 14 (1993) 812.
[11] Y.-J. Ye, J. Math. Chem., 14 (1993) 121.
[12] A. Szent-Györgyi, Nature, 148 (1941) 157; Science, 93 (1941) 609.
[13] J. Ladik, M. Seel, P. Otto and A.K. Bakhshi, Chem. Phys., 108 (1986) 203.
[14] A.K. Bakhshi, P. Otto, J. Ladik and M. Seel, Chem. Phys., 108 (1986) 215.
[15] P. Otto, A.K. Bakhshi, J. Ladik, M. Seel and S. Chin, Chem. Phys., 108 (1986) 223.
[16] A.K. Bakhshi, J. Ladik, M. Seel and P. Otto, Chem. Phys., 108 (1986) 233.
[17] A.K. Bakhshi and J. Ladik, Chem. Phys. Lett., 129 (1986) 269.
[18] P. Dean, Proc. R. Soc. London Ser. A, 254 (1960) 507; 260 (1961) 263; Rev. Mod. Phys., 44 (1972) 122.
[19] R.S. Day and F. Martino, Chem. Phys. Lett., 84 (1981) 86.
[20] Y.-J. Ye and J. Ladik, J. Math. Chem., 14 (1993) 141.
[21] A. White, P. Handler, Z.L. Smith, R.L. Hill and I.R. Lehman, Principles of Biochemistry, McGraw-Hill, New York, 6th edn., 1978, p. 78.
[22] Y.-J. Ye and J. Ladik, Phys. Rev. B, 48 (1993) 5120.
[23] Y.-J. Ye and J. Ladik, Phys. Rev. B, 51 (1995) 13091.
[24] J.F. Cutfield, S.M. Cutfield, E.J. Dodson, G.G. Donson, S.F. Emdin and C.D. Reynolds, J. Mol. Biol., 132 (1979) 85.
[25] B. Gazdy, M. Seel and J. Ladik, Chem. Phys., 86 (1984) 41.
[26] Y.-J. Ye and J. Ladik, Int. J. Quantum Chem., 52 (1994) 491.
[27] T. Odagaki and M. Lax, Phys. Rev. B, 26 (1982) 6480.
[28] H. Scher and M. Lax, Phys. Rev. B, 7 (1973) 4491.
[29] T. Odagaki and M. Lax, Phys. Rev. B, 24 (1981) 5284.
[30] J.W. Haus and K.W. Kehr, Phys. Rep., 150 (1987) 263.
[31] S. Suhai, Phys. Rev. B, 27 (1983) 3506.
[32] J. Ladik, A. Sutjianto and P. Otto, J. Mol. Struct. (Theochem), 228 (1991) 271.
[33] M. Ataka and S. Tanaka, Biopolymers, 19 (1980) 619.
[34] S. Clark and L.C. Harrison, Biochem. J., 229 (1985) 513.