

## The simulated dynamics of the insulin monomer and their relationship to the molecule's structure

P. Krüger<sup>1</sup>, W. Straßburger<sup>\*1</sup>, A. Wollmer<sup>1</sup>, W. F. van Gunsteren<sup>2</sup>, and G. G. Dodson<sup>3</sup>

<sup>1</sup> Lehr- und Forschungsgebiet Struktur und Funktion der Proteine, Abteilung Physiologische Chemie, Rheinisch-Westfälische Technische Hochschule Aachen, Klinikum Pauwelsstrasse, D-5100 Aachen, Federal Republic of Germany

<sup>2</sup> Laboratory of Physical Chemistry, The University of Groningen, Nijenborgh 16, NL-9747 AG Groningen, The Netherlands

<sup>3</sup> Department of Chemistry, University of York, Heslington, York, YO1 5DD, UK

Received June 2, 1986/Accepted in revised form January 23, 1987

**Abstract.** Insulin crystallizes in different forms, some of which show different conformations for the different molecules in the asymmetric unit. This observation leads to the question as to which conformation the molecule will adopt in solution. Molecular dynamics computer simulations of rhombohedral 2 Zn pig insulin have been carried out for both monomers (1 and 2) independently in order to study their behaviour in the absence of quaternary structure and crystal packing forces.

These preliminary 120 ps simulations suggest that both monomers converge in solution to very similar conformations which differ from the X-ray structures of both monomer 1 and 2 (Chinese nomenclature), but are closer to the former, as has previously been suggested by an analysis of the crystal packing (Chothia et al. 1983) and by energy minimization (Wodak et al. 1984). The secondary structure of the molecules is basically preserved, as expected. A detailed description of the conformational changes is given.

**Key words:** Insulin, conformers, molecular dynamics

### Introduction

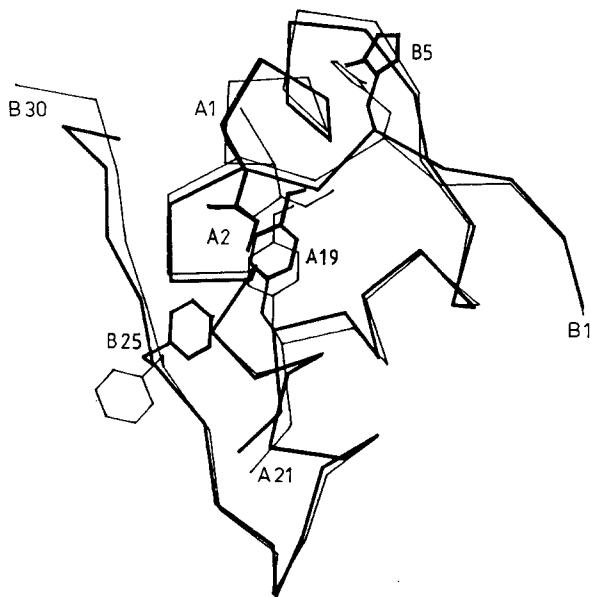
We were first confronted with the molecular dynamics of insulin when we failed to measure a large difference between the tyrosyl circular dichroism of des-Phe<sub>B1</sub>-insulin and the native hormone which was predicted by perturbation theoretical calcula-

**Abbreviations:** MD – Molecular dynamics (simulations); EM – Energy minimization; X1 – X-ray structure of molecule 1; X2 – X-ray structure of molecule 2; MD1 – mean simulated structure of molecule 1; MD2 – mean simulated structure of molecule 2; *A<sub>N</sub>* helix – helix at the *N*-terminal part of the insulin *A*-chain; *A<sub>C</sub>* helix – helix at the *C*-terminal part of the insulin *A*-chain. One and three-letter abbreviations for amino acids are used.

tions (Wollmer et al. 1977). The explanation was that the removal of Phe<sub>B1</sub> leads to increased mobility of Tyr<sub>A14</sub> (Smith et al. 1982) and, consequently, to the cancellation of its circular dichroic contributions. If Tyr<sub>A14</sub>, which is the main interaction partner of Phe<sub>B1</sub>, is also excluded from the calculations as a perturbed as well as a perturbing group the same tyrosyl CD is indeed obtained for insulin and des-Phe<sub>B1</sub>-insulin (Wollmer et al. 1980; Mercola and Wollmer 1981). The match of the observed and calculated circular dichroism can serve as a criterion of the correspondence of the structure in the crystal and in solution.

In a number of crystal forms insulin adopts more than one, usually two different conformations. The two molecules are referred to as 1 and 2 in the quaternary structure. In the 2 Zn insulin hexamer, for example, a dimer of molecule 1 and molecule 2 constitutes the asymmetric unit. Molecule 1 (Chinese nomenclature, see Cutfield et al. 1981) is the conformation present in most crystal forms and species variants whereas that of the other molecule, molecule 2, varies from one crystal field to another. These observations made on hexameric, dimeric, and monomeric insulin led to the conclusion that molecule 1 was the preferred conformation and a conformation identical with or (at least) closely similar to that of the biologically active monomer in solution.

It has been outlined (Chothia et al. 1983) how the crystal packing forces in 2 Zn insulin might induce three out of six equivalent molecules to undergo a structural transformation. The process is believed to originate at the interhexamer contact involving the end of the *A<sub>N</sub>* helix and residue His<sub>B5</sub> of molecule 2. To accommodate hexamer packing along the threefold axis a shift of the *A<sub>N</sub>* helix of molecule 2 (of an adjacent hexamer) is needed. This is however, resisted by the *A6–11* and *A7–B7* cystines and other contacts. Thus the *N*-terminal



**Fig. 1.** Superposition of molecule 1 (—) and 2 (—) from the X-ray structure of 2Zn insulin. Main chains are represented by line-connected  $C_\alpha$  atoms. Selected side chains are shown for which the orientations differ most

segment 1–6 is rotated separately ( $\Delta\phi A6 = 37^\circ$ ). This alters the position of IleA2 which makes the side chain of TyrA19 rotate and in turn displace that of PheB25. The salt bridge between A1 and B30 is also broken. Figure 1 shows where the two structures differ most.

Energy minimization with the isolated molecules 1 and 2 abolishes the main differences between them. Molecule 2 is seen to adopt the features of molecule 1 (Wodak et al. 1984). However, energy minimization procedures do not let the molecule leave a local minimum. In molecular dynamics simulations the available kinetic energy allows for transitions over energy barriers of the order of  $kT$ . The fluctuations obtained, however, can have amplitudes large enough to move the structure considerably from the X-ray structure, the starting point of the simulation.

We were interested to establish the fate of the isolated molecules structured as 1 and 2 in the course of independent dynamics simulations. Would the simulations identify “stable” conformations or sets of alternative conformations? Would any “stable” or alternative conformation be similar to one or the other of the starting geometries or different from both of them? Would the conformations converge to a common structure distinct from the starting geometries or close to one of them?

It cannot be overemphasized that the results of molecular dynamics simulations are by nature approximative. Many of the shortcomings arise from limitations in computing time. The absence of sol-

vent water which entails the omission of full charges on the protein and on the counterions, certainly is the most serious of them. Any analysis of the effects of various time-saving simplifications as, for instance, the united atom approach (van Gunsteren and Karplus 1982) would require independent comparative simulations. In the case of insulin, however, compatibility with a large body of detailed experimental information is a valuable criterion for the relevance of simulated results. Furthermore the simulations were conceived as a comparative study of molecules 1 and 2 rather than as an absolute description of insulin’s dynamics.

## Materials and methods

The insulin monomers consist of 486 atoms each. Hydrogen atoms attached to carbon atoms are incorporated into the latter, forming united atoms (van Gunsteren and Karplus 1982), whereas the other hydrogen atoms, which may form hydrogen bonds, are explicitly treated. The potential energy functions and the parameters for these functions are identical to the ones used in previous studies (van Gunsteren et al. 1983; Krüger et al. 1985), except for two changes: (i) Since the insulin simulations are performed without any discrete solvent, we have neutralized the charged atom groups (Lys, N-termini, Glu, C-termini) in order to compensate for the missing shielding effect of the solvent, as has been described by Åqvist et al. (1985). (ii) The (smaller) van der Waals parameters of van Gunsteren and Karplus (1982) are used for united atoms that are separated by three covalent bonds (third neighbours). This has the following reason: The (larger) van der Waals radii of Berendsen et al. (1984) which were used in previous studies (van Gunsteren et al. 1983; Krüger et al. 1985) do reproduce the correct crystal densities of small molecules when tested by applying molecular dynamics simulations at constant pressure (van Gunsteren and Berendsen 1985). The contraction of a protein in vacuo as observed for insulin (Wodak et al. 1984) is avoided as well. However, when used for third neighbour atoms, they induce too large a repulsion in gauche conformations. When evaluating the non-bonded interactions we have adopted a twin-range method: the total non-bonded interactions are evaluated every MD time step of  $t = 2$  fs within a short cut-off range  $R(c1) = 7\text{ \AA}$ , while the longer range (Coulomb) interactions within  $R(c2) = 20\text{ \AA}$  are evaluated less frequently, only every 10 time steps and are kept fixed between updates. We note that the cut-off radii  $R(c1)$  and  $R(c2)$  are both applied to centres of geometry of neutral atom groups

in the protein. No switching functions are used to smooth cut-off effects.

As in the previous simulations (van Gunsteren et al. 1983; Krüger et al. 1985), all bond lengths are kept rigid during the simulation by using the SHAKE method (van Gunsteren and Berendsen 1977).

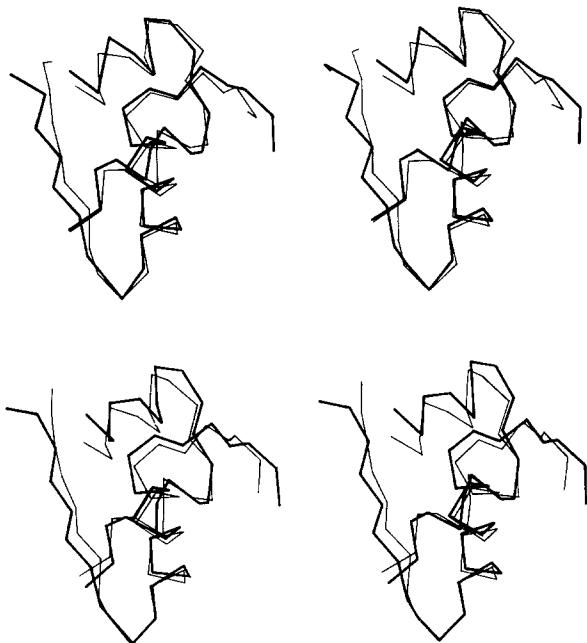
The atomic coordinates for the X-ray structure of rhombohedral 2 Zn pig insulin (Dodson, personal communication) provided the starting geometries. The initial structures of molecule 1 and 2 were not energy minimized, since they were devoid of considerable local strain. Initial velocities for the atoms were taken from a Maxwellian distribution at 300 K. The molecules were weakly coupled to the thermal bath of  $T_0 = 300$  K when integrating the equations of motion by applying the algorithm of Berendsen et al. (1984) with temperature relaxation time  $\tau = 0.1$  ps. Molecules 1 and 2 were simulated independently with the same parameters over a time span of 120 ps in vacuum. After 10 ps equilibrium is attained, therefore the analyses were performed over all transient structures from 20–120 ps at intervals of 0.1 ps. All simulations and analyses were carried out with the GROMOS library (van Gunsteren 1986). For graphic representation an interactive program called ACAMOD (Straßburger 1986) was used.

## Results and discussion

Molecular dynamics simulations are normally preceded by an energy minimization of the structure that is used as starting geometry. When carried out with the potential function of the GROMOS library, energy minimization did not result in substantial alterations of the X-ray structure. The average positional differences between the structure before and after the conjugate gradients energy minimization (EM) were of the order of 0.2 Å. This low value reflects the fact that when applying EM one moves basically only downhill over the energy hypersurface to a local minimum energy conformation, which is very similar to the initial one. Here, we did not observe the conformational change reported by Wodak et al. (1984) upon energy minimizing insulin with a different potential function. This illustrates one of the difficulties encountered with the interpretation of EM studies. Suppose two different conformations are separated by energy barriers that are lower than  $kT/2$  but significantly larger than zero. An EM study would classify the two conformations as distinct since it cannot surmount significant energy barriers, whereas in a MD simulation the molecule would be able to shuttle between the two conformations.

## General comparison of molecules 1 and 2

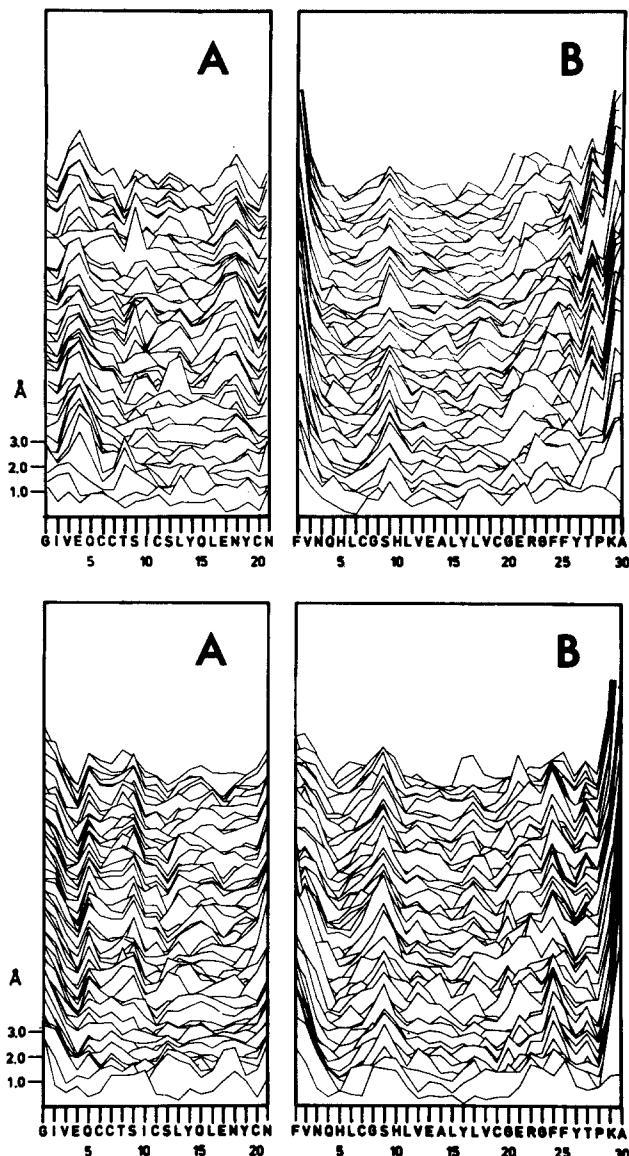
The positional differences are smallest between the averaged simulated molecules 1 and 2 (Table 1). This reflects the high degree of overlap of the conformational spaces explored in the two simulations. With a rms deviation of 0.98 Å the averaged simulated structures are closer to one another than are the two independent X-ray structures which have a deviation of 1.18 Å. They are also closer together than they are to their respective (initial) X-ray structures from which they deviate by 1.62 and 1.61 Å, respectively. The averaged simulated structure of molecule 2 is closer to the X-ray structure of molecule 1 than to that of molecule 2. The X-ray and the mean simulated structures of molecules 1 and 2 are superimposed in Fig. 2. A detailed insight into the development of the positional differences



**Fig. 2.** Superposition of the  $C_\alpha$ -atoms of the starting geometries (—) and the mean simulated structures (—) in stereo. *Top:* Molecule 1; *Bottom:* Molecule 2

**Table 1.** Deviations between the X-ray and the mean simulated structures after optimum fit on all of their  $C_\alpha$  atoms. Averages were taken over all transient structures from 20 to 120 ps of the molecular dynamics (MD) simulations. Deviations are expressed in Å

Structure	X1	X2	MD1	MD2
X1	—	1.18	1.61	1.48
X2		—	1.97	1.62
MD1			—	0.98
MD2				—



**Fig. 3 A and B.** Positional differences between the  $C_\alpha$  atoms of the X-ray structure and of transient simulated structures at 2 ps intervals from 1 (lowermost) to 120 ps (uppermost curves) of the simulation. Every curve is arbitrarily shifted upwards by 0.2 Å with respect to the preceding one for clarity of presentation. *Top:* Molecule 1; *Bottom:* Molecule 2. **A:** A-chain; **B:** B-chain

with time is offered by Fig. 3. In the A-chain deviations are encountered mainly in the  $A_N$  helix. While in molecule 2 GluA4 remains close to the X-ray structure it is a residue with particularly large differences in molecule 1. In the middle of the  $A_N$  helix the two molecules deviate in their positional differences but have a peak in common for SerA9. In the  $A_C$  helix peaks develop for GluA17 and AsnA18 of molecule 1 which are not seen in molecule 2.

As to the B-chains, PheB1 of molecule 1 starts moving away from its initial position after 4 ps. The molecules both seem surprisingly stable around

HisB5 and are displaced at SerB9, the first residue of their central  $\alpha$ -helix. Deviations are visible again at the B-chain C-terminus. While in molecule 2 both LysB29 and AlaB30 undergo a concerted motion away from their positions in the X-ray structure, B29 moves even farther away than B30 in molecule 1. A local maximum of positional differences is developed at position B25 in molecule 1 and at B24 in molecule 2.

On the whole there are no major structure deformations discernable in Fig. 3 except for the ones just mentioned which mainly occur at the very beginning of the simulation.

#### The peptide backbone

**A-chain.** Representation of the peptide backbone includes the  $N$ ,  $C$ ,  $O$ ,  $C_\alpha$  and also  $C_\beta$  atoms. The chains were subdivided into segments selected as by Wodak et al. (1984) in order to allow for differentiation in their movements. For the A-chain of molecule 1 and 2 the positional differences averaged over the segments 1–9, 10–14 and 12–19 are depicted in Fig. 4. The  $A_N$  helix can be distinguished from the other two segments by its larger positional differences in both molecule 1 and 2. Molecule 1 starts drifting away from its initial structure a little later than molecule 2. In molecule 2 segments 10–14 and 12–19, after immediate movement away from the initial structure at the very beginning, have a slight tendency to return in the course of the simulation.

There is no explicit H-bond term in the interaction potential function used here, but H-bonds are mimicked by a combination of Lennard-Jones and electrostatic terms. Therefore, it does make sense to consider the H-bonds, particularly those within secondary structural elements such as the helices. A H-bond was assumed to exist if the proton-acceptor distance did not exceed 2.5 Å and the donor-proton-acceptor angle was larger than 135°. The H-bonds in the  $A_N$ - and  $A_C$ -helices were calculated at time intervals of 1 ps. Their stability is expressed as the time fraction they were found to exist (see Table 2). The most stable bonds existing for more than 75% of the time are found at the beginning of the  $A_N$  and  $A_C$ -helices. While there is no  $\pi$ -helix type SerA9–GluA4 H-bond in the X-ray structure of molecule 2 it does exist during 48% of the simulation. In molecule 1 a corresponding result is obtained for an  $\alpha$ -helix H-bond between SerA9 and GlnA5 which is found during 55% of the simulation time; by contrast the  $\pi$ -helix type H-bond of the X-ray structure is realized only during 4% of the time. The atomic displacements which allow for the formation of these H-bonds are readily visible in

**Table 2.** H-bonds of the *A*-chain backbone. Criteria: H–O distance  $\leq 2.5 \text{ \AA}$ ; N–H–O angle  $\geq 135^\circ$ . Figures indicate how often an H-bond was found to exist at every full ps between the 20th and 120th ps of the simulation. Signs indicate that an H-bond does exist (+) or not exist (–) in the X-ray structure

Partner residues		Molecule 1		Molecule 2	
Donor [N]	Acceptor [O]	MD	X-ray	MD	X-ray
5 Gln	1 Gly	85	+	75	+
6 Cys	2 Ile	97	+	91	+
7 Cys	3 Val	80	+	91	–
8 Thr	4 Glu	10	–	2	+
8 Thr	5 Gln	31	–	5	+
9 Ser	4 Glu	4	+	48	–
9 Ser	5 Gln	55	+	12	+
16 Leu	12 Ser	77	+	80	+
17 Glu	13 Leu	90	–	97	+
18 Asn	15 Gln	10	+	8	+
19 Tyr	16 Leu	44	+	53	+
20 Cys	17 Glu	18	+	16	+

Fig. 3. While for molecule 2 a peak in the positional differences is developed at residues 5 and 9 it is observed at residues 4 and 9 for molecule 1.

Helix  $A_C$  has a somewhat higher stability in the sense that the simulations reproduce the X-ray structure better. The H-bond from residue 18 to 15 is labile. During about 17% of the simulation time there is another H-bond between Cys $A20$  and Glu $A17$ . The  $\alpha$ -helical parts of the H-bond pattern ( $i+4 \rightarrow i$ ) are well reproduced in the simulation, whereas the  $3_{10}$ -helical parts ( $i+3 \rightarrow i$ ) show less correspondence with the X-ray structure (see Table 2). Differences in H-bond pattern between molecules 1 and 2 point to a decreased H-bond stability. So, it is not surprising that for these H-bonds differences between X-ray and simulation results are observed.

It has been suggested that a  $37^\circ$  rotation of the  $\varphi A6$  angle could make the  $A_N$  region in molecule 2 adopt the conformation of that in molecule 1 (Chothia et al. 1983). The initial values for this angle are  $-101^\circ$  in the X-ray structure of molecule 1 and  $-62^\circ$  in that of molecule 2. During simulation the  $\varphi A6$  angles of both molecules vary within amplitudes of typically  $40^\circ$  between these values but there is no pronounced reorientation towards one or other of them. The calculations suggest this helical fragment is structurally flexible.

*B*-chain. the *B*-chain was subdivided into segments: *B*1–*B*9, *B*9–*B*19, *B*20–*B*24, and *B*22–*B*29. In Fig. 5 the positional differences from the initial structure (for averages over the *N*,  $C_\alpha$ ,  $C_\beta$ , *C* and *O* atoms) are plotted segmentwise as a function of time. Deviations from the X-ray structure are more

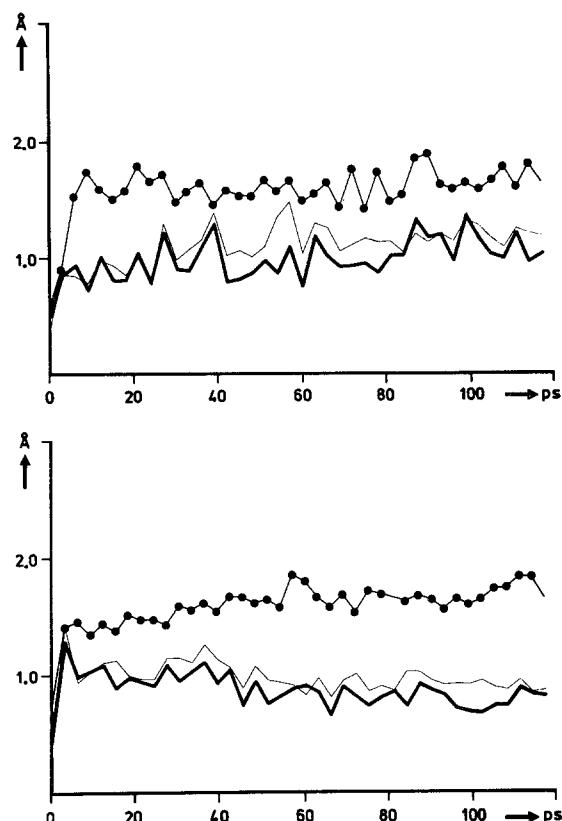
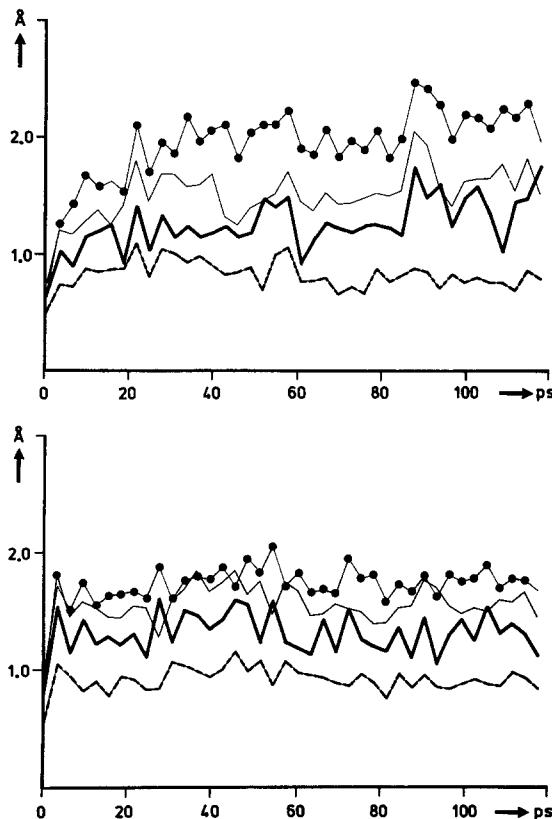


Fig. 4. Positional differences between the main chain atoms (*N*,  $C_\alpha$ ,  $C_\beta$ , *C*, *O*) of the X-ray structure and of transient simulated structures at 1 ps intervals as a function of simulation time. Averages over the segments: ●, A1–A9; □, A10–A14; ▲, A12–A19. Top: Molecule 1; Bottom: Molecule 2

pronounced for *B*22–*B*29 than for the other segments in both molecule 1 and 2 and are largest in the former. Compared with the *C*-terminus the *N*-terminal segment is surprisingly stable, except for Phe*B*1 which has fairly large positional differences particularly in molecule 1 (see Fig. 3). This may be an outcome of the cystine bridge which possibly keeps the residues neighbouring Cys*B*7 in place. The molecules do not differ much with respect to their *B*-chain *N*-termini. The difference in the mobility of Phe*B*1 seems not to exert a significant influence on the first segment as a whole. The readiness of the latter to adopt drastically different arrangements in des[pentapeptide *B*26–30]insulin and 4 Zn insulin cannot be anticipated from the calculations. The stability of the central *B*-helix is also reflected in its H-bond pattern (see Table 3). In the simulations there is a certain tendency for the helix to be adjusted to  $\alpha$ -type bonding and extended towards the *N*-terminus. While in the X-ray structure Leu*B*11 is consistently H-bonded to Gly*B*8 it is preferentially paired with Cys*B*7 during 66% of

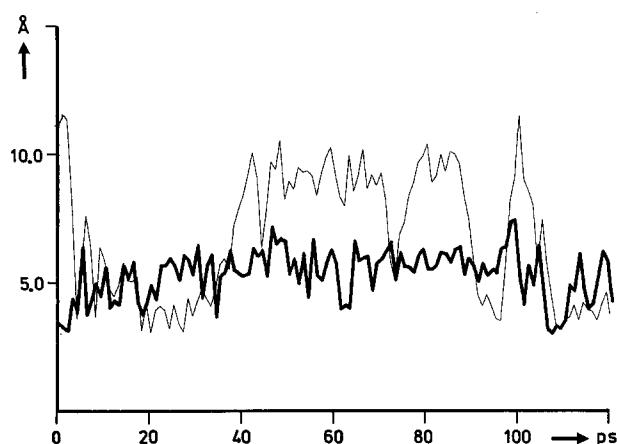


**Fig. 5.** Positional differences between the main chain atoms ( $N$ ,  $C_\alpha$ ,  $C_\beta$ ,  $C$ ,  $O$ ) of the X-ray structure and of transient simulated structures at 1 ps intervals as a function of simulation time. Averages over the segments: —,  $B1-B9$ ; - - -,  $B9-B19$ ; - · - ,  $B20-B24$ ; · · · · · ,  $B22-B29$ . Top: Molecule 1; Bottom: Molecule 2

the simulation time, while Gly  $B8$  clearly is the partner of Val  $B12$  in both molecule 1 and 2. In the X-ray structure of molecule 1 H-bond criteria are met simultaneously for Leu  $B11 \rightarrow$  Gly  $B8$  and Val  $B12 \rightarrow$  Gly  $B8$ . During the dynamics simulations there is a clear move towards  $\alpha$ -helical pairing.

#### Interhelical distances

In order to detect movements of the helices relative to each other they were fitted by corresponding stretches of ideal  $\alpha$ -helix (Lesk and Hardman 1982) and the direction cosine of the axial vectors of the helices (Chou et al. 1983) was calculated as a function of simulation time. The angles between the vectors for any pair of helices ( $A_N/A_C$ ;  $A_N/B$ ;  $A_C/B$ ) typically vary by  $\pm 9^\circ$  about a constant value without showing any uniform trend (not depicted). The same applies to the distances between the geometric centres of any pair of helices (not depicted either). Thus, rigid body movements upon relaxation of the crystal and hexamer constraints as sug-



**Fig. 6.** Distance between the  $N$ -terminus of the  $A$ -chain and the  $C$ -terminus of the  $B$ -chain as a function of simulation time. The distance was calculated between the mass centres of the protons of the  $A1$  amino group and of the oxygens of the  $B30$  carboxylate group. —, molecule 1; - - -, molecule 2

**Table 3.** H-bonds of the  $B$ -chain backbone. Criteria:  $H-O$  distance  $\leq 2.5$  Å;  $N-H-O$  angle  $\geq 135^\circ$ . Figures indicate how often an H-bond was found to exist at every full ps between the 20th and 120th ps of the simulation. Signs indicate that an H-bond does exist (+) or not exist (-) in the X-ray structure

Donor [N]	Acceptor [O]	Partner residues		Molecule 1		Molecule 2	
		MD	X-ray	MD	X-ray	MD	X-ray
11 Leu	7 Cys	66	-	40	-		
11 Leu	8 Gly	4	+	14	+		
12 Val	8 Gly	94	+	92	-		
13 Glu	9 Ser	79	+	89	+		
14 Ala	10 His	99	+	95	+		
15 Leu	11 Leu	98	+	97	+		
16 Tyr	12 Val	95	+	90	+		
17 Leu	13 Glu	93	+	87	+		
18 Val	14 Ala	91	+	87	+		

gested by Chothia et al. (1983) are not detectable with the criteria used here.

*Distances of chain termini.* In the X-ray structure a hydrogen of the  $\alpha$ -ammonium group of Gly  $A1$  and an oxygen of the terminal carboxylate group of Ala  $B30$  are at salt bridge distance in molecule 1 whereas they are 11 Å apart in molecule 2. Figure 6 shows how these distances vary in the course of the simulations. While in molecule 1 the "salt bridge" is always maintained, in molecule 2 the partner groups approach, separate and reapproach, thus the "salt bridge" alternating between being formed and broken. Although for Lys  $B29$  the differences are similar in molecule 1 and 2, the last residue is tied down in molecule 1 only. It seems possible that the

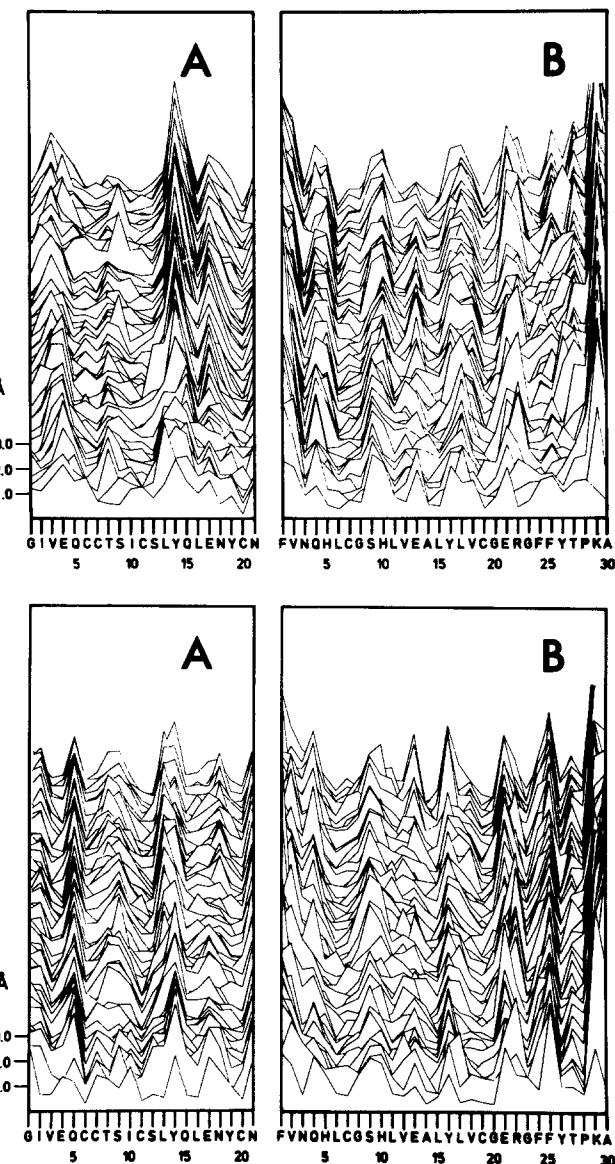
deviations in the  $A_N$  helices contribute to the differences in stability. This interpretation is supported by the finding that residues  $A1$  and  $A2$  of molecule 1 differ less from the initial structure than those of molecule 2 (compare in Fig. 3).

#### Side chains

**General view.** Deviations in the side chains were obtained in a calculation analogous to that for the main chain atoms. The average positional differences for the side chain atoms of every residue at every ps of the simulation are plotted in Fig. 7. As is to be expected there are relatively high contributions from long side chains with high intrinsic flexibility, such as those of lysine or glutamine. The contribution of aromatic side chains depends on their environment, e.g. on whether they are buried in the interior of the molecule or exposed at its surface. They will be dealt with below. Figure 7 is much in line with the corresponding plot for the main chain atoms (Fig. 3). This applies to the large deviations for GlnA5 in monomer 2 and for the *N*- and *C*-terminus of the *B*-chain as well as to the shift of high deviation from *B*25 to *B*24 in molecule 2 and from *B*24 to *B*25 in molecule 1.

However, the behaviour of the side chains can also differ from that of the main chain, for instance at *A*14 Tyr where the appearance of a peak after 35 ps is restricted to the side-chain atoms in molecule 1. With the information displayed in Figs. 3 and 7 it is possible to identify the residues which can move and make a conformational change. The X-ray structures of the molecules differ with respect to an H-bond between the hydroxyl group of Tyr*A*19 and the  $O_e$  of Gln*A*5 lacking in molecule 2 — though reported to form with energy minimization (Wodak et al. 1984). In the simulation the distance between the hydroxyl and the amide group immediately shrinks from 9 to about 3.5 Å in molecule 2 and then remains constant. This corresponds to a peak at position *A*5 in Fig. 7 which builds up at the very beginning of the simulation.

**Focus on the aromatic side chains.** Aromatic side chains are involved in one of the most prominent differences between molecules 1 and 2 in the X-ray structure (see Fig. 1). Phe*B*25 of molecule 2 is oriented towards the neighbouring monomer over the antiparallel  $\beta$ -sheet, while in molecule 1 it is folded back onto the molecule it is part of. The question arises as to whether in the course of the dynamics simulation one of the side chains would adopt the orientation of the other. In molecule 2  $\chi_1$  of Phe*B*25, which is  $-50^\circ$  in the X-ray struc-



**Fig. 7A and B.** Positional differences between the averaged side-chain atoms of the X-ray structures at 2 ps intervals from 1 (lowermost) to 120 ps (uppermost curves) of the simulation. Every curve is arbitrarily shifted upwards by 0.2 Å with respect to the preceding one for clarity of presentation. Top: Molecule 1; Bottom: Molecule 2. **A:** *A*-chain; **B:** *B*-chain

ture, initially drops to a value of about  $-160^\circ$  where it remains except for a very short return to its original value after 60 ps. The behaviour of  $\chi_1$  is very different in molecule 1: Starting at  $+100^\circ$ , its value in the X-ray structure, it adopts three different plateaux, each for roughly 40 ps. From the first plateau at  $+60^\circ$   $\chi_1$  drops within a few ps to  $-160^\circ$ , i.e. the same plateau as obtained in the simulation for molecule 2. The third plateau roughly corresponds to the value of  $\chi_1$  in the X-ray structure of molecule 2. A single dihedral angle, however, owing to correlated changes of others, will not necessarily provide a full and unambiguous reflection of a

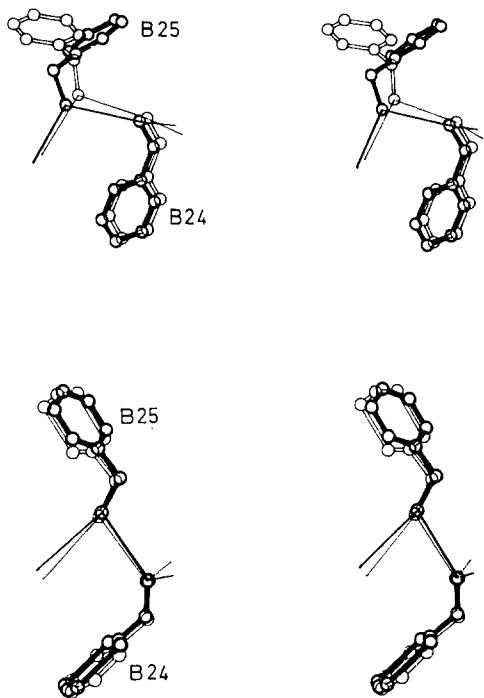


Fig. 8. Stereo drawing of the relative positions of Phe B24 and Phe B25 in the X-ray structure (—) and in the simulated structure, time-averaged over ps 35–75 (—). Top: Molecule 1; Bottom: Molecule 2

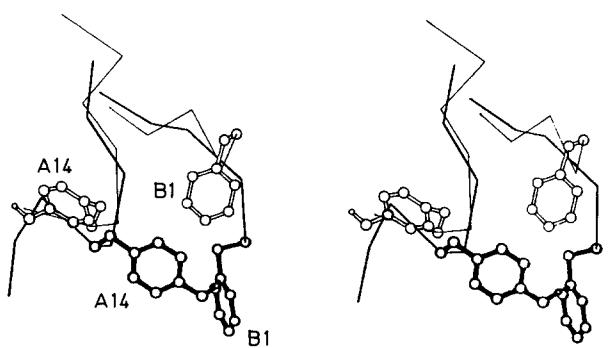


Fig. 9. Stereo drawing of the relative positions of Tyr A14 and Phe B1 of molecule 1 in the X-ray structure (—) and in the simulated structure, time-averaged over ps 35–75 (—)

structural transition (McCammon and Karplus 1979; McCammon et al. 1983). Thus, after the first re-orientation of  $\chi_1$  in molecule 1 the  $C_y - C_z$  axis through the phenyl ring no longer points towards the molecule (see Fig. 8). In molecule 2 there is a concerted movement of the Phe B24 ring outwards and the Phe B25 ring inwards with the  $\chi_1$  angle remaining constant. This withdraws the Phe B25 ring from above the  $\beta$ -sheet. On the whole the Phe B25 rings in both molecules converge to positions which are very similar but different from those in the X-ray structure of both molecule 1 and 2.

The stability of  $\chi_1$  in molecule 2 after its initial rotation suggests transition into an energetically

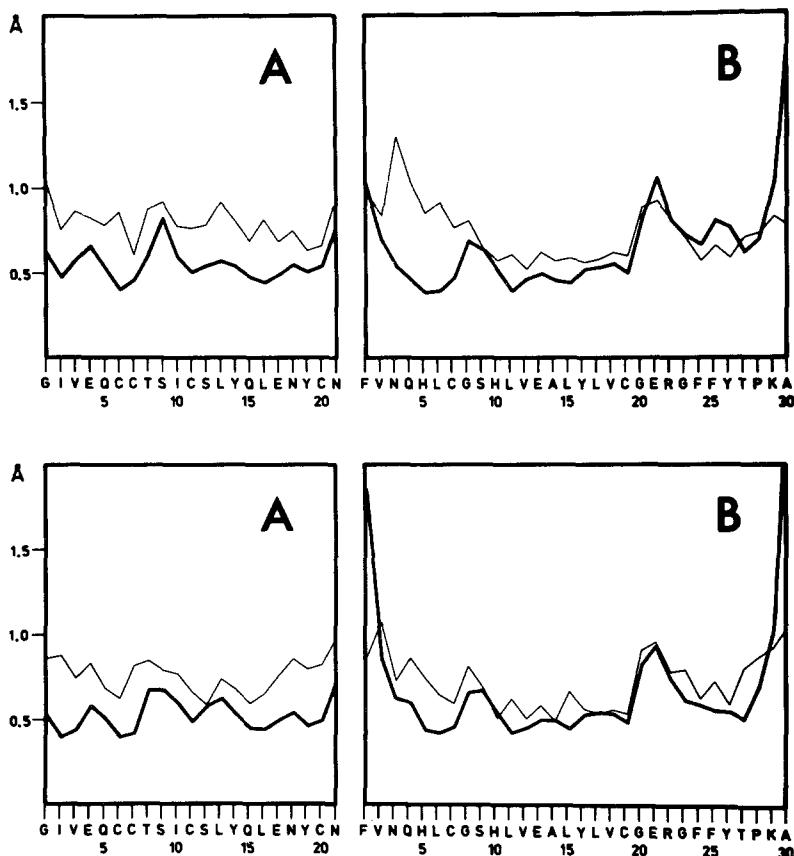
preferred conformation. The three plateaux accessible for  $\chi_1$  in molecule 1, on the other hand, may represent energetically similar alternative orientations.

Tyr A19 appears very stable. Therefore the distances calculated between the ring centres of Tyr A19 and the other aromatic side chains should provide a measure of any relative movements. The A19–A14 distance in molecule 1 shows a sudden decrease of this distance from 13 to 10 Å after 30 ps, which seems to be related to the appearance of a peak in the positional differences for Tyr A14 (compare top of Fig. 7). The shift of Tyr A14 is linked to a displacement of Phe B1. Their relative positions after 30 ps of simulation and in the X-ray structure, where they are forming a hydrophobic surface involved in hexamerization, are depicted in Fig. 9 for molecule 1. An analogous behaviour is not observed for molecule 2. Although the N-terminus of its B-chain is also very flexible a corresponding shift of the Tyr A14 side chain does not occur within the duration of the simulation.

#### Comparison of dynamics as obtained in the simulations and reflected in experiment

One possibility of comparing simulation to experiment is to compare rms fluctuations derived from the simulation and atomic temperature factors obtained in the X-ray analysis and refinement. Even simulations neglecting the solvent and crystal force fields tend to give an acceptable correspondence with the B-values, since for a globular protein the amplitudes of atomic motions are largely determined by the internal forces.

The rms fluctuations for the  $C_\alpha$  atoms of molecule 1 and 2 as derived from the B-values of the X-ray structure ( $\text{rms} = (3B/8\pi^2)^{1/2}$ ) and from the time interval 20–120 ps of the simulations are compared in Fig. 10. Their general pattern is very similar though the B-value derived fluctuations are consistently higher than the simulated ones. This may arise from static and crystal disorder and, possibly, the rather short duration of the simulation. There is a poorer match in the A-chain where differences of typically 0.4 Å are encountered. In the A-chain the B-values tend to be higher than in the B-chain which is physically sensible since the A-chain lies largely on the hexamer surface whereas the B-chain is mostly buried. The B-derived motion is much less than that calculated for Phe B1 of molecule 2. This is in contrast to the positional differences which are larger for molecule 1 than for molecule 2 (see Fig. 3). Apparently Phe B1 of molecule 1 has moved and adopted a more stable position



**Fig. 10 A and B.** Rms fluctuations as calculated from 20 to 120 ps of the dynamics simulation (—) and from the atomic *B*-values of the X-ray analysis (—). Comparison for the  $C_{\alpha}$  atoms. *Top*: Molecule 1; *Bottom*: Molecule 2. A: *A*-chain; B: *B*-chain

before the 20th ps. Residues *B*3–*B*7 of molecule 1, however, exhibit rather large differences in mobility, possibly due to unrecognized disorder. From *B*8 onwards differences are small for both molecules, even in the highly flexible turn region which follows the central helix. The largest difference of all is found for the last residue.

The fluctuations of the side-chain atoms are not depicted. Again those derived from the simulations tend to be higher than the *B*-value derived ones. They are particularly high for polar side chains like Glu, Gln, Asn and Lys where large size and surface location combine, but they are also above average for exposed aromatic side chains. The residues Tyr*A*14, Phe*B*1, His*B*10, and Tyr*B*16, as well as Phe*B*24 and *B*25 in molecule 1 all have markedly higher motions in the simulations. This is to be expected since they are involved in subunit interactions in the 2Zn insulin crystal. According to Chothia et al. (1983) the constraints causing the different conformers to form involve the *A*<sub>N</sub> helix, mainly of molecule 2, and His*B*5 of molecule 1. However, it seems incompatible with this interpretation that the difference between the simulation and the *B*-value derived fluctuations is the same for His*B*5 in molecule 1 and 2. Furthermore the distances between his*B*5 (centre of  $C_{\gamma}$ – $N_{\epsilon}$  connection)

and the geometric centre of the *A*<sub>N</sub> helix ( $C_{\alpha}$  atoms) in molecule 1 and 2 do not change in the course of the simulation (not depicted).

On the whole the correspondence in the patterns of simulation and *B*-value derived rms fluctuations is also remarkable for the side chains. The smaller amplitudes of the *B*-value derived fluctuations is a reflection of the general damping effect exerted by solvent, quaternary structure and packing contacts in the crystal.

### Conclusion

Important objectives in simulation of molecular dynamics are to differentiate a protein with respect to mobility potential and, related to this, identify onsets for possible conformational changes or structural pathways between known conformations. Molecular calculations may also help to clarify the effects on a protein's structure produced by aggregation and crystal contacts. These calculations on insulin have been carried out on a monomer with conformations 1 and 2, corresponding to molecules 1 and 2 in the 2Zn insulin asymmetric unit. There is good evidence that molecule 1 is nearer the con-

formation that occurs in solution. It is encouraging that the calculations show that molecule 2 moves to new conformations for the most part significantly nearer that of molecule 1, in a generally understandable pattern. The large positional changes at *B*1 and the *B*-chain *C*-terminus are at residues which are affected by hexamer and crystal packing; being on the monomer surface they are therefore less constrained and can be expected to move. There is a distinct movement at the *A*-chain *N*-terminus of molecule 2 again understandable as adjustments from the perturbations introduced by crystal packing (Chothia et al. 1983). There are local movements at Ser *B*9 and Glu *A*4 which probably derive from relaxation of the hexameric organisation and crystal contacts, respectively.

The calculations reveal possible patterns of structural behaviour in the molecule. The *A*-chain *N*-terminal helix for example has two distinct conformations in molecules 1 and 2. During the simulations molecule 1 and 2 are found to form and reform the H-bonds characteristic of the alternative conformations. This suggests that the alterations seen in insulin molecule 2, which permit close crystal packing, are easily assumed by the unconstrained molecule. By contrast the central stretch of helix between *B*11 and *B*18 is very stable. The two H-bonds possible at *B*11 NH *B*8 CO have a very similar geometry in the crystal structure; the simulations suggest the  $\alpha$ -helical ( $i+4 \rightarrow i$ ) contacts are preferred. It is possible that this preference is associated with the conformational changes already noted at Ser *B*9.

The changes occurring at *B*24 and *B*25 are unexpected. It is possible that they are produced by the absence of water in the simulations. This deficiency may also be responsible for the apparent conformational instability in the *A*-chain *N*-terminus, which like *B*24 and *B*25 in the monomer, is able to H-bond to water molecules. In a less dramatic fashion the H-bonding to water by *B*10 and *B*13 carbonyl *O* may also stabilise the helical contacts which appear a little variable in these calculations. The conformational behaviour of *B*24 and *B*25 in these simulations justifies further research since it may be biologically significant. The observation by Kobayashi et al. (1982) shows that D-Phe at *B*24 elevates potency. This substitution must perturb the folding of insulin from its crystal structure. It appears therefore that there are structural changes produced by D-Phe *B*24, and that these favour insulin's receptor binding and action. A tendency for this kind of structural change to occur intrinsically in the native molecule suggested by these simulations may be relevant to insulin's action. The fact that the calculations were carried out in the absence

of water may in a way correspond to the presumed non-polar environment of the receptor.

The simulation calculations on insulin have got to be further developed if they are to guide us in analysing the hormone's behaviour. The indications concerning the molecule's conformational flexibility are important in understanding how insulin moves between its different crystal structures and how perhaps these structures relate to biological activity. Further research will be directed towards including water molecules as well as selected insulin derivatives and establishing their influence on insulin's conformational behaviour.

**Acknowledgements.** The Aachen group gratefully acknowledges valuable discussions with Dr. S. J. Wodak and Madame M. Prévost, Bruxelles. Independently, these colleagues continue simulating the dynamics of insulin in vacuum using different potential functions and shorter time steps as well as accounting for every proton explicitly. After detailed analyses a comparison of the two simulations will be communicated in a joint publication. The skillful technical assistance of Mrs. Angelika Szameit is gratefully acknowledged. Thanks are due to the Abteilung Medizinische Statistik und Dokumentation for the use of the VAX 11/780 computer and to the Rechenzentrum der RWTH Aachen for access to the microfilm unit. A 10-min colour movie was prepared showing the molecular dynamics of the insulin molecule.

Our work was supported by grants from the Deutsche Forschungsgemeinschaft (Wo 152/8-4 and Wo 152/7-2) and the Boehringer Ingelheim Fonds.

## References

- Aqvist J, Gunsteren WF van, Leijonmarck M, Tapia O (1985) A molecular dynamics study of the *C*-terminal fragment of the L7/L12 ribosomal protein. *J Mol Biol* 183:461-477
- Berendsen HJC, Postma JPM, Gunsteren WF van, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* 81:3684-3690
- Chothia C, Lesk AM, Dodson GG, Hodgkin DC (1983) Transmission of conformational change in insulin. *Nature* 302:500-505
- Chou K-C, Nemethy G, Scheraga HA (1983) Energetic approach to the packing of  $\alpha$ -helices. I. Equivalent helices. *J Phys Chem* 87:2869-2881
- Cutfield JF, Cutfield SM, Dodson EJ, Dodson GG, Reynolds CD, Vallely D (1981) Similarities and differences in the crystal structures of insulin. In: Dodson G, Glusker JP, Sayre D (eds) *Structural studies on molecules of biological interest - A volume in honour of Dorothy Hodgkin*. Clarendon Press, Oxford, pp 527-546
- Gunsteren WF van (1986) GROMOS Program System. Distributed through BIOMOS biomolecular software b.v. Laboratory of Physical Chemistry, University of Groningen, NL
- Gunsteren WF van, Berendsen HJC (1977) Algorithms for macromolecular dynamics and constrained dynamics. *Mol Phys* 34:1311-1327
- Gunsteren WF van, Berendsen HJC (1985) Molecular dynamics simulations: Techniques and applications to proteins. In: Hermans J (ed) *Molecular dynamics and protein structure*. Polycrystal Book Service, Western Springs, Illinois, USA, pp 5-14

Gunsteren WF van, Karplus M (1982) Effect of constraints on the dynamics of macromolecules. *Macromolecules* 15: 1528–1544

Gunsteren WF van, Berendsen HJC, Hermans J, Hol WGI, Postma JPM (1983) Computer simulation of the dynamics of hydrated protein crystals and its comparison with X-ray data. *Proc Natl Acad Sci USA* 80: 4315–4319

Kobayashi M, Ohgaku S, Iwasaki M, Maegawa H, Shigeta Y, Inouye K (1982) Supernormal insulin: D-Phe B24 insulin with increased affinity for insulin receptors. *Biochem Biophys Res Commun* 107: 329–336

Krüger P, Straßburger W, Wollmer A, Gunsteren WF van (1985) A comparison of the structure and dynamics of avian pancreatic polypeptide hormone in solution and in the crystal. *Eur Biophys J* 13: 77–88

Lesk AM, Hardman KD (1982) Computer-generated schematic diagrams of protein structures. *Science* 216: 539–540

McCammon JA, Karplus M (1979) Dynamics of activated processes in globular proteins. *Proc Natl Acad Sci USA* 76: 3585–3589

McCammon JA, Lee CY, Northrup SH (1983) Side-chain rotational isomerization in proteins: A mechanism involving gating and transient packing defects. *J Am Chem Soc* 105: 2232–2237

Mercola D, Wollmer A (1981) The crystal structure of insulin and solution phenomena: use of the high-resolution structure in the calculation of the optical activity of the tyrosyl residues. In: Dodson G, Glusker JP, Sayre D (eds) *Structural studies on molecules of biological interest – A volume in honour of Dorothy Hodgkin*. Clarendon Press, Oxford, pp 557–582

Smith GD, Duax WL, Dodson EJ, Dodson GG, de Graaf RAG, Reynolds CD (1982) The structure of des-Phe B1 bovine insulin. *Acta Crystallogr B* 38: 3028–3032

Straßburger W (1986) *Molekulare Dynamik von Proteinen*. Habilitationsschrift, RWTH Aachen

Wodak SJ, Alard P, Delhaise P, Renneboog-Squilbin C (1984) Simulation of conformational changes in 2Zn insulin. *J Mol Biol* 181: 317–322

Wollmer A, Fleischhauer J, Straßburger W, Thiele H, Brandenburg D, Dodson G, Mercola D (1977) Side-chain mobility and the calculation of tyrosyl circular dichroism of proteins. Implications of a test with insulin and des-B 1-phenylalanine insulin. *Biophys J* 20: 233–243

Wollmer A, Straßburger W, Hoenjet E, Glatter U, Fleischhauer J, Mercola DA, de Graaf RAG, Dodson EJ, Dodson GG, Smith DG, Brandenburg D, Danho W (1980) Correlation of structural details of insulin in the crystal and in solution. In: Brandenburg D, Wollmer A (eds) *Insulin – chemistry, structure and function of insulin and related hormones*. Walter de Gruyter, Berlin New York, pp 27–35