NMR solution spatial structure of 'short' scorpion insectotoxin I_5A

Alexander S. Arseniev, Vladimir I. Kondakov, Vladimir N. Maiorov and Vladimir F. Bystrov*

M.M. Shemyakin Institute of Bioorganic Chemistry, USSR Academy of Sciences, Ul. Vavilova-32, 117988, GSP-1, Moscow V-334, USSR

Received 13 October 1983

Two-dimensional 500 MHz NMR study reveals the three-dimensional structure of the insectotoxin I_5A of *Buthus eupeus* in aqueous solution. The most likely set of disulfide linkages is proposed. Comparison with the single crystal structure of the 'long' toxin v-3 of *Centruroides sculpturatus* shows similarity in their α-helical and antiparallel β-structure fragments.

Buthus eupeus *insectotoxin I_5A* *NMR* *Nuclear Overhauser effect* *Peptide conformation* *Scorpion toxin*

1. INTRODUCTION

Scorpion polypeptide neurotoxins are characterized by high selectivity in their biological action – the venom of one scorpion species contains toxic compounds acting on nervous systems only of vertebrates or of arthropods [1]. The homologous family of the 'long' neurotoxins (60–70 amino acid residues) is relatively well studied and even the crystal structure of the toxin v-3 of *Centruroides sculpturatus* Ewing is established [2]. The 'short' insectotoxins received less attention and there is almost no information on their spatial structure. These toxins with 35–36 amino acid residues and 4 disulfide bonds, the localization of which is not yet known, manifest a paralysing effect on insects, presumably by acting on the glutamate receptor of the postsynaptic membrane [3].

This paper deals with the NMR conformational analysis of the *Buthus eupeus* short polypeptide insectotoxin I_5A in solution. The most feasible set of the S-S bridges is also proposed. The exposed spatial structure is compared with the X-ray result on the long toxin v-3 [2].

* To whom correspondence should be addressed

2. MATERIALS AND METHODS

Insectotoxin I_5A isolated from *B. eupeus* venom [3] was kindly provided by Dr E.V. Grishin. Initially the NMR spectra of the compound were analyzed on the basis of the insectotoxin I_5 *B. eupeus* primary structure [3]. However, in the process of signal assignment some structural discrepancies were delineated and later confirmed by direct sequencing [4]. The Lys, Asn and Gly residues were identified in positions 14, 23 and 24 of the insectotoxin I_5A primary structure (fig.1) in-

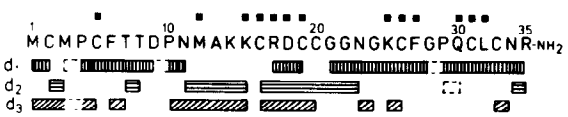

Fig.1. Primary structure of *B. eupeus* insectotoxin I_5A [4] and survey of the sequential NOE connectivities d_1–d_3: d_1, between the $C_i^{\alpha}H$ and $N_{i+1}H$ protons; d_2, between N_iH and $N_{i+1}H$; d_3, between $C_i^{\beta}H$ and $N_{i+1}H$. The fragments indicated by dashed lines correspond to the specific case of proline residues for which the NOE interactions of the Pro^{i+1} $C^{\delta}H_2$ protons with protons of the preceding residue are observed. The residues with slowly exchanged peptide NH protons are marked by solid squares above the sequence.

Published by Elsevier Science Publishers B.V.
00145793/84/$3.00

stead of the respective Asn, Gly and Lys in insectotoxin I_5.

The two-dimensional (2-D) NMR spectra COSY, SECSY and NOESY [5] were recorded with a 500 MHz Bruker WM-500 spectrometer as in [4]. The 11 mM solutions of insectotoxin I_5A in H_2O (with 10% ^2H_2O) and in ^2H_2O were placed in 5 mm NMR tubes and studied at pH 2.9 and 5.5 and at 30 and 50°C.

The distance geometry algorithm computer software was generously provided by Dr W. Braun. The program [6] allows estimation of the coordinates of N ($\leqslant 150$) atoms consistent with the distance constraints

$$L_{ij} \leqslant |r_i - r_j| \leqslant U_{ij}$$

where L_{ij} and U_{ij} denote, respectively, the lower and upper limits on the distance $|r_i - r_j|$ between the atomic centers i and j. We used the pseudoatomic description of the molecule, i.e., each of the amino acid residues was represented by two spherical pseudoatoms α for the backbone NH-C^{α}H-CO fragment and β for the side chain.

3. RESULTS AND DISCUSSION

By combined use of the COSY, SECSY and NOESY spectra practically all the proton signals of insectotoxin I_5A were assigned according to the primary structure [4]. The NOESY spectra reveal interproton direct dipole–dipole interactions demonstrated by the nuclear Overhauser effect (NOE) [7]. The NOE between protons of the particular amino acid residue in combination with the indirect spin–spin proton coupling constants HN-C^{α}H and HC^{α}-C^{β}H describes the local conformation (torsional angles ϕ and χ^1) of the residue. The NOE between protons of neighboring residues, in particular the d_1, d_2 and d_3 interactions depicted in fig.1, are used for signal assignment [4], to eliminate some ambiguity in the ϕ and χ^1 angles, and for ψ angle estimation. Additionally, in the NOESY spectra of insectotoxin I_5A more than 100 NOE interactions were detected between protons located in non-neighboring amino acid residues which while remote in the primary structure become close to one another owing to folding of the peptide backbone [7].

The conformational analysis of insectotoxin I_5A proceeds in 3 steps – configuration of the X–Pro amide bonds, secondary structure and three-dimensional structure.

There are 3 proline residues in fragments Met^3–Pro^4, Asp^9–Pro^{10} and Gly^{28}–Pro^{29} of the insectotoxin molecule. From general structural considerations it is obvious that the distances from the $C^{\alpha}H$ proton of the i-th residue to the $C^{\alpha}H$ and $C^{\delta}H_2$ protons of the proline $i+1$ residue depend on the *trans* or *cis* configuration of the X^i–Pro^{i+1} peptide bond and on the ψ angle of the i-th residue. Calculations with the standard geometry of amino acid residues [8] give the following distances, respectively, for *trans* and *cis* X–Pro bonds: H_i^{α}–H_{i+1}^{α} 0.43–0.48 nm and 0.18–0.38 nm; H_i^{α}–H_{i+1}^{δ} 0.20–0.39 nm and 0.43–0.50 nm. Taking into account that under the employed NOESY experimental conditions (mixing time 100 ms) we observe NOE cross-peaks for protons separated by less than 0.30–0.35 nm [4], it is clear that for the *trans* X–Pro bond we expect only $H_i^{\alpha}...H_{i+1}^{\delta}$ dipolar interaction. Indeed, only NOE cross-peaks for the $C_i^{\alpha}H$ and $C_{i+1}^{\delta}H_2$ protons are observed, but no $C_i^{\alpha}H...C_{i+1}^{\alpha}$ peaks inherent to the *cis*-configuration are detected. Thus this proves that all 3 X–Pro bonds in insectotoxin I_5A are in the *trans* configuration.

The secondary structure evaluation is based on consideration of extreme values of d_1, d_2 and d_3 distances for principal components of the regular protein secondary structure (table 1). Calculation was performed by changing in the region of ± 20° the standard values of the ϕ and ψ torsional angles taken correspondingly as −56°, −44° for right-handed α-helix; −114°, 114° for parallel and −136°, 134° for antiparallel β-structures. In the d_3 calculation the χ^1 angle was allowed to rotate freely from −180° to +180°. From table 1 it follows

Table 1

Extreme values of d_1, d_2 and d_3 distances (in nm) in the peptide fragment for regular polypeptide secondary structure components – right hand α-helix (α_r), parallel and antiparallel β-structures (β_p, β_a)

Distance	α_r	β_p	β_a
d_1	0.35–0.36[a]	0.22–0.23	0.22–0.24
d_2	0.24–0.33	0.38–0.45	0.29–0.45
d_3	0.21–0.41	0.32–0.46	0.27–0.44

[a] For the glycine residue 0.26–0.30 nm

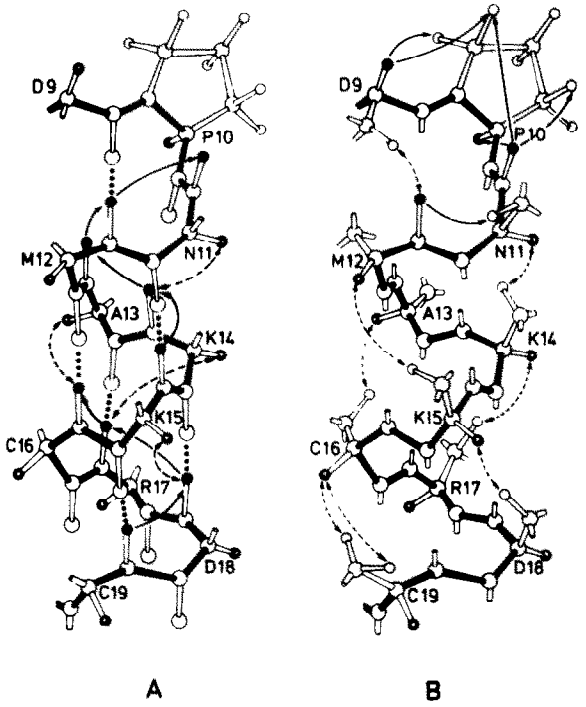

Fig.2. Right hand α-helical conformation of the Asp[9]–Cys[19] fragment of *B. eupeus* insectotoxin I_5A molecule. Arrows indicate interproton NOE interactions (A) between the backbone NH and $C^{\alpha}H$ protons and (B) between the backbone and side chain protons. Solid and dashed arrows correspond to the neighboring residue (fig.1) and to the non-neighboring residue proton dipolar interaction, respectively. By rows of 4 points are shown the interresidue hydrogen bonds – for participating NH groups very slow exchange with deuterium was observed in D_2O solution ($t_{1/2}$ from 38 to >100 h at pH 2.7 and 32°C).

that for α-helical parts of the polypeptide chain we expect to observe the sequential NOE d_2 cross-peaks without any d_1 peaks (except of glycine residues). In contrast, the presence of the d_1 peaks and absence of d_2 interactions is inherent to β-structures. The d_3 peaks normally should be absent in the case of the parallel β-structure.

From the observed set of *d* interactions (fig.1) it follows that the right-hand α-helical part extends from the Asn[11] to at least the Cys[19] residue of insectotoxin I_5A. The β-structure is accomplished in the Asn[23]–Phe[27] and Gln[30]–Asn[34] fragments. Because these two fragments are located nearby in the sequence one could assume that they are arranged in the antiparallel β-structure with the β-bend constituted by Gly[28] and Pro[29] residues. The CD and laser Raman spectra are consistent with right-hand α-helix and β-structure content of the insectotoxin I_5A secondary structure. The partial results of the NMR conformational analysis are shown in fig.2,3.

Confirmation of the proposed regular secondary structure and construction of three-dimensional structure for corresponding fragments of insectotoxin I_5A are based on examination of additional NMR data of hydrogen-bonded backbone amide NH groups, torsional angles estimated from the vicinal proton couplings, and non-neighboring residue interproton NOE interactions [4,7]. For instance, the residues located in the α-helical fragment have $^3J(H\text{-}NC^{\alpha}\text{-}H)$ values in the range of 2.4–5.0 Hz, which corresponds to ϕ angles of $-70 \pm 10°$ [9]. The ambiguity of the $^3J(H\text{-}NC^{\alpha}\text{-}H)$

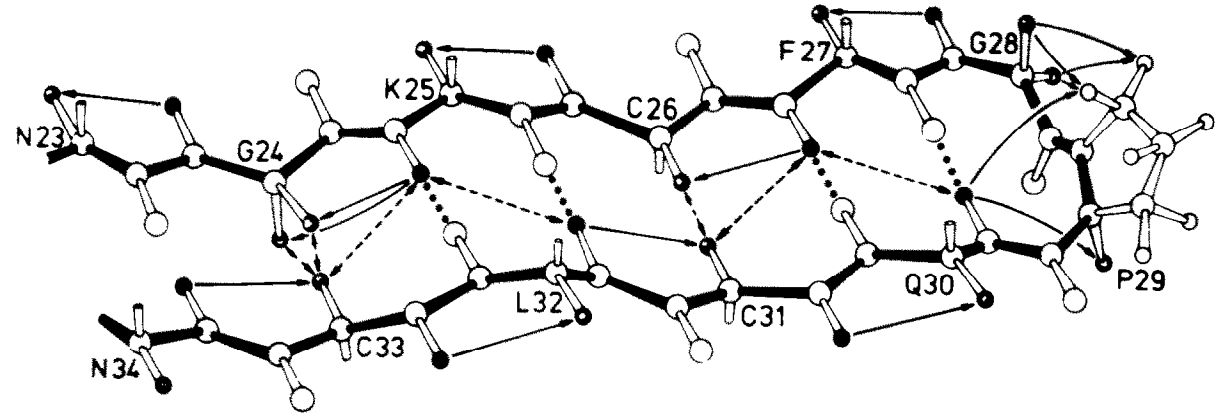

Fig.3. Conformation of antiparallel β-structure in the Asn[23]–Asn[34] fragment of *B. eupeus* insectotoxin I_5A. For further explanation see fig.2 legend; $t_{1/2}$ > 100 h.

and ϕ angle relationship could be removed by taking into account the proton NOE results on intraresidue and interresidue d_2 distances. The ^3J(H-NC$^\alpha$-H) coupling values 7.0–10.1 Hz for the 25–27 and 30–33 residues are in agreement with the β-structure ϕ angles $-140 \pm 15°$ [9].

The overall spatial structure of insectotoxin I_5A was determined using the distance geometry algorithm in pseudoatomic approximation. The NOE connectivities between protons of non-neighboring residues (fig.4) were employed as distance constraints – if protons from the two pseudoatoms demonstrated the cross-peak in the NOESY spectrum it was assumed that the distance between the surfaces of these two spherical pseudoatoms was less than 0.15 nm. Additionally, the packing density of polypeptide molecules was taken into account [10] and in the quadratic error function was included the corresponding term:

$$F(r_1,\ldots,r_N, R_o) = [R^2(r_1,\ldots,r_N) - R_o^2]^2,$$

where $R(r_1,\ldots,r_N)$ is the radius of gyration evaluated from coordinates of pseudoatoms [11] and R_o is the target value of the radius estimated from the protein packing density and the total volume of atomic groups assembling the molecule taking into account the nonspherical shape of the molecule [10]. From the distance geometry algorithm 15 dense packed spatial structures of insectotoxin I_5A obeying the distance constraints (fig.4) were obtained with the error function in the range 10^{-7}–10^{-12} nm^4. The comparison by the procedure in [11] shows that an average root-mean-square deviation between these structures is 0.21 ± 0.03 nm (for only α-pseudoatoms 0.17 ± 0.03 nm), i.e., they are conformationally close to one another.

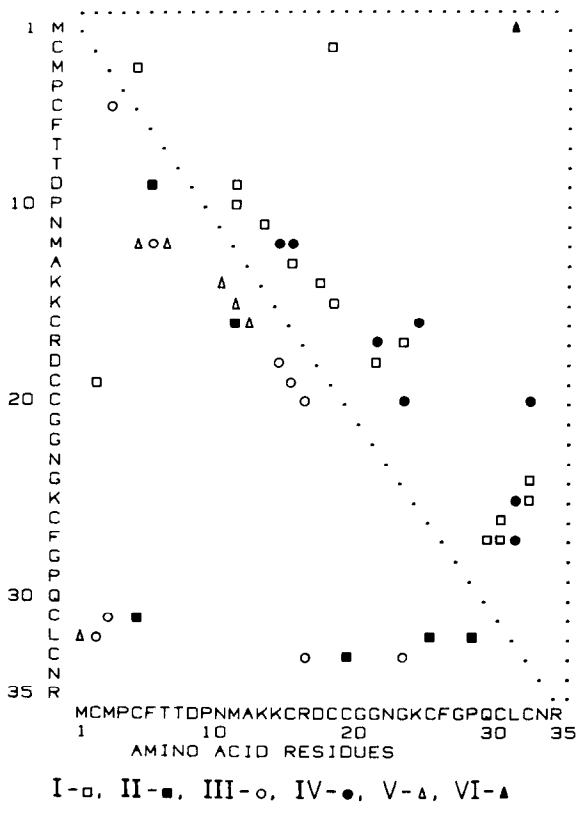

Fig.4. Map of contacts between the α- and β-pseudoatoms of non-neighboring residues of insectotoxin I_5A: I, $\alpha\alpha$; II, $\beta\beta$; III, $\alpha\beta$; IV, $\beta\alpha$; V, $\beta\alpha$ and $\alpha\beta$; VI, $\alpha\alpha$ and $\beta\alpha$. Each contact is shown only once above or below the diagonal line.

The secondary structure analysis of the α-pseudoatomic coordinates in the 15 structures of insectotoxin I_5A by the method in [12] clearly reveals helical fragment Asn^{11}–Cys^{20} and two fragments with the extended conformation (Asn^{23}–Phe^{27} and Gln^{30}–Asn^{34}) connected by the reverse turn Phe^{27}–Gln^{30}. The tendency to form the reverse turn is also displayed in the connection fragment Cys^{20}–Asn^{23} between the helical and extended structures as well as in the N-terminal part (Asp^9–Met^{12}) of the helix. The rest of the N-terminal fragment does not tend to form the regular secondary structure. Thus the secondary structure sections determined directly from the NMR data on local conformation (fig.2,3) and from the overall spatial structure generated by the distance geometry algorithm in pseudoatomic approximation completely agree.

Unfortunately in the conformational NMR analysis of insectotoxin I_5A use could not be made of the 4 disulfide bonds because they have not yet been localized by the chemical technique [3]. As far as we are aware, there are no other direct procedures for establishing the S–S linkages. Thus we attempted to solve this problem by the statistical analysis of feasible disulfide bridges in the obtained 15 spatial structures of insectotoxin I_5A taking into account restriction rules based on the secondary structure of the molecule [13].

Since all 8 cystine residues in the 15 spatial structures are grouped relatively close to one another,

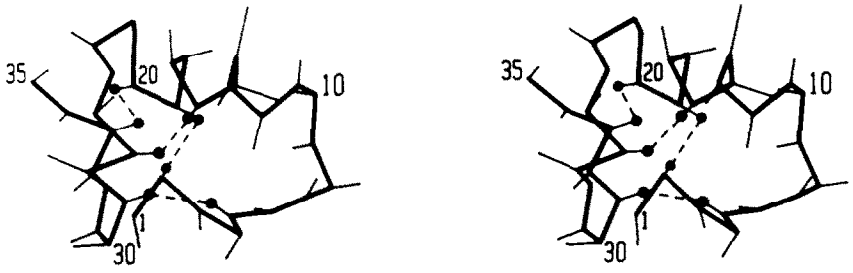

Fig.5. Computer stereoscopic presentation of the *B. eupeus* insectotoxin I_5A conformation, obtained by distance geometry algorithm on the basis of NMR data. Filled circles represent the side chains of cysteine residues.

as a result only 10 sets of 4 S–S bonds were found to be possible [7]. However, one set looks more realistic because the β-pseudoatoms of the 3 pairs of Cys residues (2–19, 5–31 and 20–33) are situated at the shortest distances in all 15 structures and if so then the fourth S–S bridge has to be formed by Cys^{16} and Cys^{26}. A stereoscopic view of the insectotoxin I_5A backbone three-dimensional structure with these disulfide bonds pattern is shown in fig.5.

In conclusion it is worthwhile to compare the obtained solution structure of short insectotoxin I_5A with the recently established X-ray crystal structure of the long type scorpion toxin v-3 (65 amino acid residues) of *C. sculpturatus* Ewing [2]. In spite of the absence of any homology in amino acid sequences and different biological action [1–3,14], the spatial structures of these scorpion toxins contain many closely related facets (fig.6). Both have α-helical and antiparallel β-structural fragments with nearly the same number of amino acid residues. Moreover these fragments are similarly packed to one another in space. However, toxin v-3 has an additional 30 residues located in the N- and C-terminal parts of the molecule. It is assumed [2] that exactly these parts

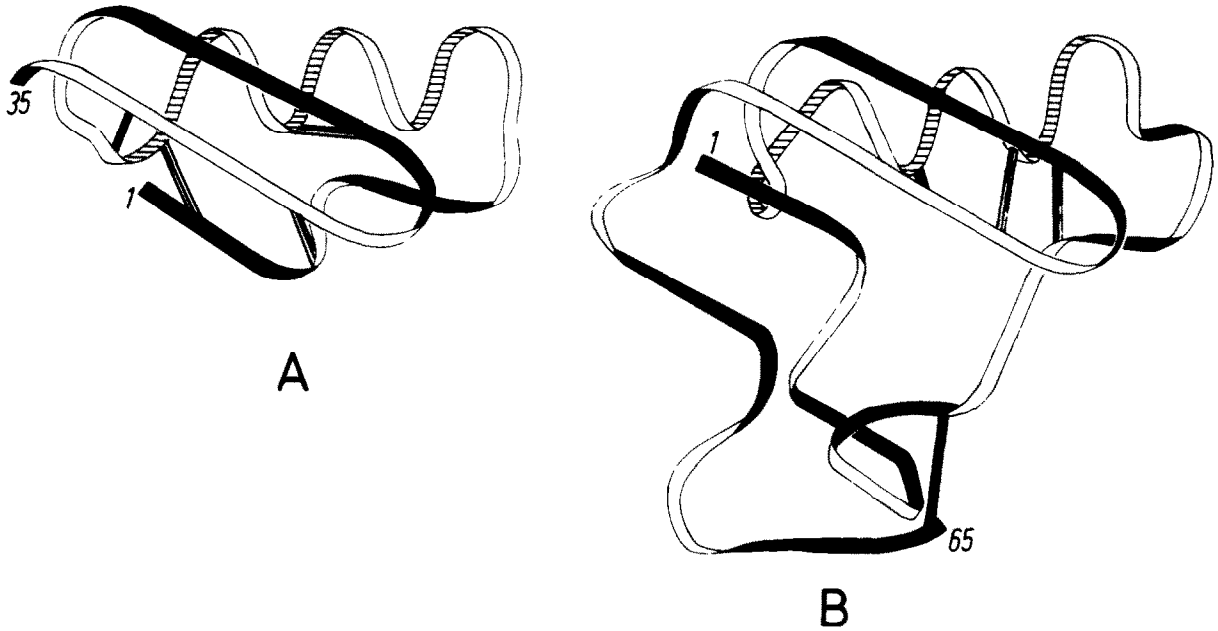

Fig.6. Schematic comparison of backbone folding of short and long types of scorpion toxins. (A) Solution structure of *B. eupeus* insectotoxin I_5A. (B) X-ray crystal structure of *C. sculpturatus* Ewing toxin v-3, as redrawn from [2].

of long toxins are responsible for their biological action. Thus, it is likely that the structural similarity of short and long scorpion toxins is not connected at least directly with their mechanism of action. Presumably these types of scorpion toxins possess the same predecessor but are at different stages of evolution.

REFERENCES

[1] Grishin, E.V., Soldatov, N.M., Tashmuchamedov, B.A. and Atakuziev, B.U. (1978) Bioorgan. Khem. (USSR) 4, 450–461.
[2] Fontecilla-Camps, J.C., Almassy, R.J., Suddath, F.L. and Bugg, C.E. (1982) Toxicon 20, 1–7.
[3] Grishin, E.V., Volkova, T.M. and Soldatova, L.N. (1982) Bioorgan. Khem. (USSR) 8, 155–164.
[4] Arseniev, A.S., Kondakov, V.I., Maiorov, V.N., Volkova, T.M., Grishin, E.V., Bystrov, V.F. and Ovchinnikov, Yu.A. (1983) Bioorg. Khem. (USSR) 9, 768–793.
[5] Bax, A. (1982) Two-Dimensional Nuclear Magnetic Resonance in Liquids, Delft University Press, London.
[6] Braun, W., Bosh, C., Brown, L.R., Go, N. and Wuthrich, K. (1981) Biochim. Biophys. Acta 667, 377–396.
[7] Arseniev, A.S., Kondakov, V.I., Maiorov, V.N., Bystrov, V.F. and Ovchinnikov, Yu.A. (1983) Bioorg. Khem. (USSR), in press.
[8] Momany, F.A., McGuire, R.F., Burgess, A.W. and Sheraga, H.A. (1975) J. Phys. Chem. 79, 2361–2381.
[9] Bystrov, V.F. (1976) Progr. NMR Spectrosc. 10, 41–81.
[10] Richards, F.M. (1974) J. Mol. Biol. 82, 1–14.
[11] Oobataka, M. and Crippen, G.M. (1981) J. Phys. Chem. 85, 1187–1197.
[12] Ramakrishnan, C. and Soman, K.V. (1982) Int. J. Peptide Protein Res. 20, 218–237.
[13] Thornton, J.M. (1981) J. Mol. Biol. 151, 261–287.
[14] Beneski, D.A. and Catteral, W.A. (1980) Proc. Natl. Acad. Sci. USA 77, 639–643.