Conservative Val^{47} residue of POU homeodomain: role in DNA recognition

Alexander G. Stepchenko*, Nadezda N. Luchina, Oleg L. Polanovsky

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov str. 32, Moscow 117984, Russia

Received 14 April 1997

Abstract Conservative Val^{47} residue, located in the third recognition helix of the Oct-2 POU domain, was alternately substituted with other 19 amino acids. Affinity and specificity of interaction with oct-site ATGCAAANGA and homeo-specific site ATAANGA were determined for all mutants. The wild type protein (with Val^{47}) has maximal affinity and specificity in POU domain interaction with octamer sequence. However, V47I mutant showed stronger interaction with homeo-specific site. The highest specificity of interaction with homeo-site was recorded for V47S mutant. We conclude that only Val^{47} provides sequence-specific high-affinity binding of POU proteins with octamer targets other than the homeo-specific site. It is shown also that damages caused by point mutations may be at least partially compensated by participation in the oct-site recognition of both POUh and POUs domains.

Key words: Oct-2 protein; POU domain mutants; Oct-sequence; Homeo-specific site; Protein-DNA recognition

1. Introduction

The POU domain proteins represent a family of proteins involved in various aspects of transcription regulation [1,2]. They possess a DNA-binding POU domain consisting of N-terminal POU-specific domain (POUs) and C-terminal POU homeodomain (POU$_H$), connected by a variable polypeptide linker. Both domains are required for high-affinity sequence-specific binding. Previous studies have shown that POUs domain is closely related to the DNA-binding domains of phage λ and 434 repressors [3,4] whereas the POU$_H$ domain is similar to classic homeodomains [5]. In DNA binding, the individual domains each use a helix-turn-helix (HTH) unit for recognition, and helix 3 from each domain is docked in the major groove of the DNA [6–8]. The optimal DNA-binding site for Oct-1 or Oct-2 proteins is an octamer sequence (5′-ATGCAAAT-3′) [9]. The recognition sequences of POUs and POU$_H$ domains correspond to the 5′-half (ATGC) and 3′-half (AAAT) of the POU domain-binding site, respectively [8,9]. Additional bonds may be formed in the DNA minor groove by amino acids of N-terminal arm of the POU$_H$ domain [8]. The sequences recognized by isolated POUs and POU$_H$ domains at some positions differ from those recognized by the entire POU domain, suggesting that the POU domain-binding site is more than simple juxtaposition of the POUs and POU$_H$ target sequences [9]. The invariant amino acid residues in third recognition helix of POU$_H$ domain form contacts with the bases of the 3′ part of octamer site AAAT. Val^{47} contacts thymine at position 4 [8], Asn^{51} contacts adenine at position 3 [8,10], and Cys^{50} (according to X-ray analysis) contacts the TT base pair 3′-flanking the octamer site [8]. Although the best binding site for Oct-1 and Oct-2 proteins is the octamer sequence, they also bind with a low affinity to the various DNA targets such as TAATGARAT sequences [9]. All of these targets contain TAAT core sequence, common for homeoprotein-binding sites [8,11,12]. Three amino acid residues, Val, Ile and Asn, in position 47 of homeodomain are critical for recognition of a distal thymine of the TAAT core sequence. Val residue at the same position of the POU$_H$ domain is absolutely conserved in contrast to homeoproteins [8].

This work was aimed to study Val^{47} contribution to octamer site and homeo-specific site recognition. This residue was alternately substituted by 19 other amino acids and interaction of mutant proteins with DNA targets was studied. The results demonstrated that Val^{47} is critical for sequence-specific high-affinity POU domain binding with octamer targets, except homeo-specific sites, and this explains why it is conserved in POU proteins.

2. Materials and methods

2.1. Construction of expression plasmids and purification of proteins

The cDNA fragment (amino acids 179–392) [13] containing POU domain was cloned in pUR292 plasmid between *Bam*HI and *Hin*dIII sites [11] and expressed in *E. coli* BMH 71-18. The cells were grown to 0.3 units of OD_{600} in LB medium supplemented with 8 g/l of Brain Heart Infusion, and 125 μg/ml of ampicillin, then IPTG was added to 0.5 mM. After 2 h induction, bacterial cells were harvested by centrifugation and resuspended in 6 volumes of a buffer (10 mM Tris-HCl pH 8.0, 10 mM EDTA, 1 mM PMSF, 0.2 mg/ml of lysozyme), incubated for 20 min at 0°C, and freeze-thawed in the presence of 0.1% NP40. Bacterial debris and nucleic acids were pelleted by centrifugation. Proteins were precipitated with ammonium sulfate added to 35% saturation. The pellet, collected by centrifugation, was resuspended in PBS containing 1 mM DTT, 1 mM PMSF, and dialyzed against the same buffer. Then the protein solution was passed through the Sepharose 2B affinity column with immobilized rabbit polyclonal antibodies against total *E. coli* BMH 71-18 proteins. The flow-through was collected and dialyzed against buffer containing 100 mM KCl, 40 mM Tris-HCl, pH 7.8, 1 mM EDTA, 1 mM DTT, 1 mM PMSF, and 50% glycerol. Protein concentration and purity were monitored by densitometry of gels stained with Coomassie blue (purified wild-type POU protein was used as a standard of protein concentration).

2.2. Site-directed mutagenesis

A *Hin*dIII-*Bam*HI fragment encoding the POU domain was cloned in pTZ18U vector (Pharmacia). Phage DNA preparation and mutagenesis were performed for all possible amino acid substitutions of Val^{47} using the Pharmacia Oligonucleotide-Directed in vitro Mutagenesis System with mutagenic oligonucleotide:

5′-CGGTTGCAGAACCANNNGCGGATCACTTCCTTCTCC-3′

where N is A, G, T or C. Mutations were identified by Sanger sequencing.

*Corresponding author. Fax: (7) (95) 135 14 05.
E-mail: pol@imb.msk.su

PII S0014-5793(97)00508-5

2.3. Oligonucleotides

Two sets of 35-bp oligonucleotides ('TAAT' and 'OCT')

'TAAT' 5′-AGGTACCTGAGTTGATAANGAGACTGTCTCTAGAG-3′ (1 2 3 4 5 6 7)

'OCT' 5′-AGGTACCTGAGATGCAAANGAGACTGTCTCTAGAG-3′ (1 2 3 4 5 6 7 8 9 10)

were synthesized to assay the DNA binding to the wild-type and mutant POU proteins. Each set contained four oligonucleotides with the randomized nucleotide at positions 5 and 8 of the homeo-specific binding site TAANGA or octamer site ATGCAAANGA, respectively. Flanking sequences were identical and did not contain additional TAAT and ATGC to exclude the effect of these sequences on the probe binding.

2.4. Gel retardation assay

Binding reaction was performed with of 0.1 ng of ^{32}P-labeled oligonucleotide and variable amounts of POU protein in 20 μl of binding buffer: 20 mM Tris-HCl, pH 7.8, 50 mM KCl, 1 mM EDTA, 1 mM PMSF, 5% glycerol, 100 μg/ml of BSA, 1 mM DTT, 0.25 μg of poly[d(G-C) poly d(G-C)]. After 1 h incubation at 20°C the protein-bound and free DNA were resolved in 5% TE polyacrylamide gel (Fig. 1A). Gel slices corresponding to free and bound DNAs were cut out and quantified by counting in a scintillation liquid. Amount of free DNA was plotted vs. log of POU domain concentration (Fig. 1B). Dissociation constant was determined as the protein concentration required with excess protein for binding 50% of a DNA probe and approximated to that of the wild-type protein.

3. Results and discussion

Val^{47} of POU$_H$ domain is directly adjacent to T8 of the octamer-binding site ATGCAAAT, and one of valine methyl groups is in Van der Waals contact with thymine methyl group in DNA major groove [8]. Similar interaction was reported for Ile^{47} of *Engrailed* [14] and MATα2 homeodomains [15] with T4 of homeo-specific sequence TAAT. No stringent contacts between T4 and amino acid residues of POU$_H$ domain were detected in DNA minor groove. Thus, Val^{47} determines the position of T4 nucleotide in the core sequence TAAT or T8 in the octamer site ATGCAAAT.

To evaluate the role of Val^{47} in DNA recognition, this residue was alternatively substituted by 19 other amino acids. DNA-binding properties of mutant proteins were tested with oligonucleotides containing oct-sequence ATGCAAANGA or homeo-specific binding site ATAANTGA. Affinity of the wild-type POU domain and its mutants was estimated in the mobility-shift assay (Fig. 1A). The equilibrium dissociation constant of protein-DNA complex was determined at fixed concentration of DNA probe and increasing concentration of POU domain (Fig. 1B). As shown in Fig. 2A, when T8 of octamer-binding site ATGCAAATGA was substituted with A, C or G, the binding activity of wild-type POU was drastically decreased. Relative affinities for these altered sites were 5–14-fold less than for the wild-type octamer sequence. In contrast, all mutants recognize bases at position 8 of the octamer sequence ATGCAAANGA with more relaxed specificity than the same bases in homeo-specific site ATAANGA (Fig. 2A and B). Substitutions of Val^{47} with bulky or long-size hydrophobic Leu, Phe, Trp, Pro, or with acidic Asp and Glu result in loss of DNA binding to both octamer and homeo-specific binding sites (not shown). V47M, V47Q, V47K and V47H mutants still bound to the octamer sites but less efficiently than the wild-type POU domain (Fig. 2A). Relative affinities of these mutants to homeo-specific sites decreased significantly (Fig. 2B). Val^{47} substitution with Ala or sharply reduced the POU domain affinity, though it induced no steric

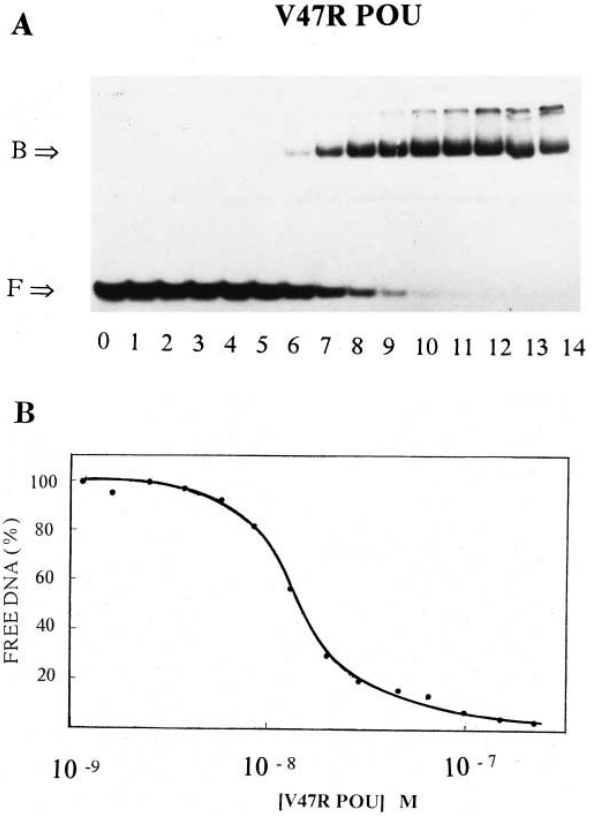

Fig. 1. Gel retardation assay of DNA-protein complexes. **(A)** Labelled oligonucleotide incubated with increasing amounts of V47R POU mutant protein. 0, control DNA probe without protein. Positions of the protein-DNA complex (B) and free DNA (F) are indicated by arrows. **(B)** Quantification of gel retardation data. Amount of free DNA as a function of log protein concentration. Relative affinity of protein to octamer probe (ATGCAAATGA) was determined as described in Section 2.

tension. This effect may be caused by the incorrect placement of the POU$_H$ domain third helix in the DNA major groove. The substitution of the conservative Val^{47} by Arg, Gln, Lys, or Cys amino acids perturbed the ability of POU domain to discriminate between different bases at the position 8 of the octamer site. Wild-type and V47I (to a lesser extent V47T and V47S mutants) demonstrated similar binding preference when T8 of octamer site has been substituted with C, A or G. Both Val^{47} and Ile^{47} form Van der Waals contacts with methyl groups of distal thymine in octamer ATGCAAAT and homeo-specific TAAT-binding sites [8,14,16]. Methyl groups of Val and Ile occupy a similar position in DNA major groove, therefore the V47I substitution induce minimal effect on the POU domain interaction with octamer site (Fig. 2A). Nevertheless, V47I and V47S mutants discriminate bases at position 5 of homeo-specific site more efficiently and bases at position 8 of octamer less efficiently than the wild-type POU. Reduced ability of the mutant proteins ability to discriminate between different bases at position 8 of the octamer site (in comparison with homologous position of homeo-site) may be partially explained by cooperative interaction of POU$_H$ and POUs domains with DNA [8,17]. We mentioned above that both POU$_H$ and POUs domains participate in oct-site recognition whereas homeo-specific sequence is recognized by POU$_H$ domain. There are no protein-protein contacts between POU$_H$

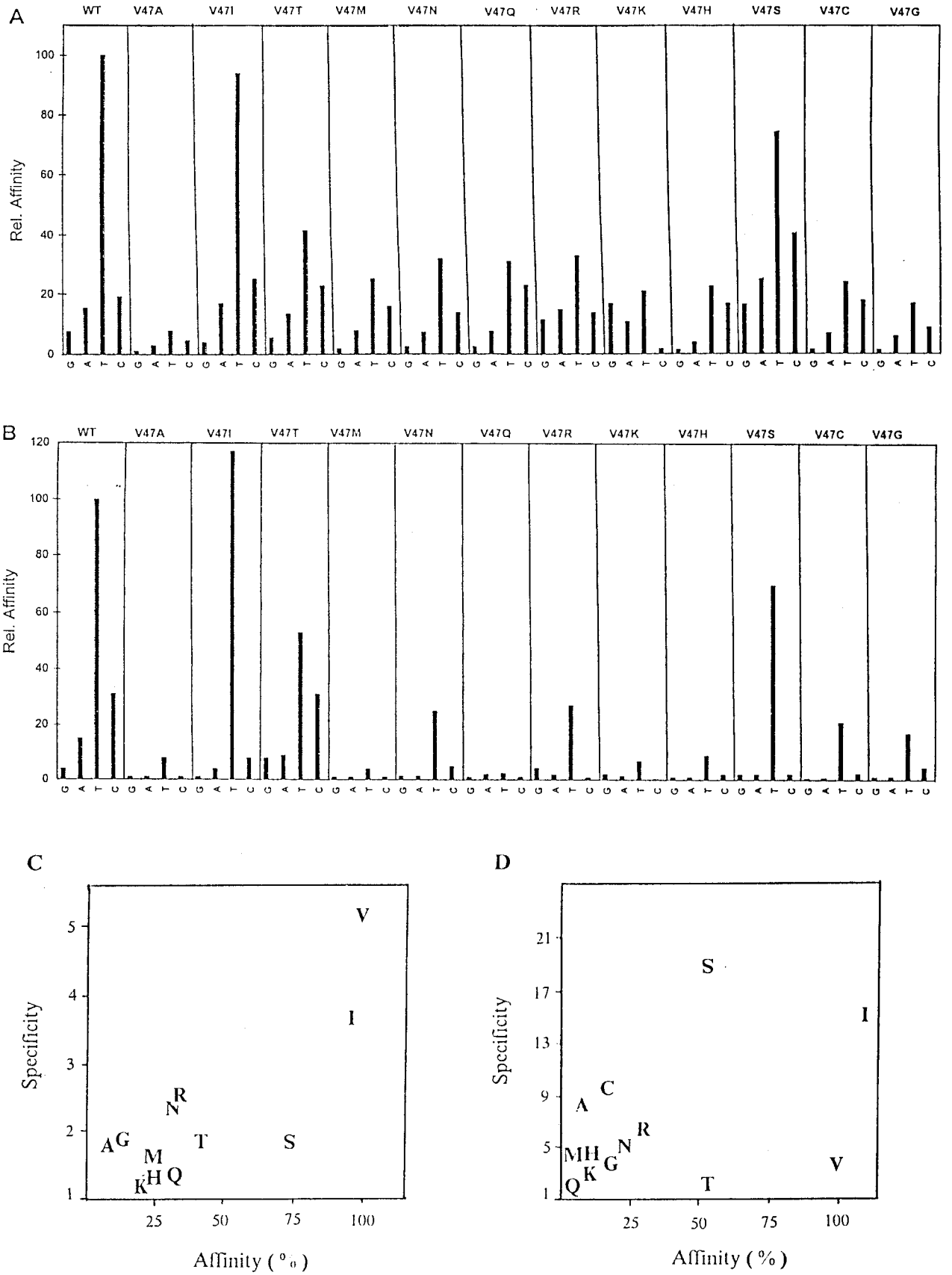

Fig. 2. Comparison of the Oct-2 POU domain affinity to octamer and homeo-specific sites. Relative affinity (Rel. Affinity) was normalized to that of the wild-type protein. **(A)** Effect of base-pair substitution (G, A, T, C) in position 8 of oct-sequence ATGCAAANGA. **(B)** The same for substitution at position 5 of homeo-specific site ATAANGA. **(C)** Specificity/affinity plots of POU mutants for oct-site. **(D)** The same for homeo-specific site. Mutant proteins are designated by the letters corresponding to amino acid residues at position 47 of POUH domain.

and POUs domains. However, one domain alters binding affinity and specificity of the other due to overlapping phosphate contacts near the center of the octamer site [8]. The indirect interaction of POU_H and POUs domains may result also from changes in the DNA-binding site conformation [18].

Concerning specificity and affinity to octamer site, the mutants may be arranged as shown on Fig. 2C. The specificity of POU domain-DNA interaction was presented as a ratio of the maximal binding site affinity to the lower affinity. Among 20 amino acids tested, only Val^{47} (wild-type protein) and partially, Ile^{47} contributed to the maximal affinity and specificity of POU domain interaction with octamer sequence other than homeo-specific sites (Fig. 2C and D). That may explain why most of POU proteins have Val at position 47 of POU_H domain. Only Pit-1 contains Ile at the same position, but its binding sites differ from the canonical oct sequence [19].

Affinity/specificity plots of the mutants for the homeo-sites are essentially different from oct-sites (Fig. 2D). Ile^{47} and to a less extent, Ser^{47} are optimal, whereas Val is much less preferable. That is in accordance with occurrence of Ile in equivalent positions of homeodomains [20].

It seems that contacts formed by the POU_H Val^{47} and Asn^{51} residues are essential for the POU_H domain third helix correct orientation in the DNA major groove. This, in turn, allows basic amino acid residues to contact with the sugar-phosphate backbone of DNA, and thus to determine the affinity of interaction. The importance and cooperativity of such unspecific contacts is confirmed by our data that substitutions of the POU homeodomain Arg^{46} by Val, His, Thr or Lys result in at least 10-fold reduction in DNA binding (Stepchenko, unpublished).

Acknowledgements: The authors are grateful to Prof. W. Schaffner and Prof. V.I. Ivanov for participation at the initial steps of this work. We thank Dr. Liza Pankratova for helpful discussion and comments. This work was supported by International Science Foundation, Russian National Human Genome Program and DOE Humanitarian Grant.

References

[1] Ruvkun, G. and Finney, M. (1991) Cell 64, 475–478.
[2] Rosenfeld, M.G. (1991) Genes Dev. 5, 897–907.
[3] Anderson, W.F., Ohlendorf, D.H., Takeda, Y. and Matthews, B.W. (1981) Nature 298, 754–758.
[4] McKay, D.B. and Steitz, T.A. (1981) Nature 290, 744–749.
[5] Pabo, C.O. and Sauer, R.T. (1992) Annu. Rev. Biochem. 61, 1053–1095.
[6] Assa-Munt, N., Mortshire-Smith, R.J., Aurora, R., Herr, W. and Wright, P.E. (1993) Cell 73, 193–205.
[7] Dekker, N., Cox, M., Boelens, R., Verrijzer, C.P., Van der Vliet, P.C. and Kaptein, R. (1993) Nature 362, 852–855.
[8] Klemm, J.D., Rould, M.A., Aurora, R., Herr, W. and Pabo, C.O. (1994) Cell 77, 21–32.
[9] Verrijzer, C.P., Alkema, M.J., Weperen, W.W., Van Leeuwen, H.C., Strating, M.J.J. and Van der Vliet, P.C. (1992) EMBO J. 11, 4993–5003.
[10] Botfield, M.C., Jansco, A. and Weiss, M.A. (1994) Biochemistry 33, 8113–8121.
[11] Stepchenko, A.G. (1992) Dokl. Ross. Akad. Nauk (Moscow) 335, 175–179.
[12] Stepchenko, A.G. (1994) FEBS Lett. 337, 175–178.
[13] Clerc, R.G., Corcoran, L.M., LeBowitz, J.H., Baltimore, D. and Sharp, P.A. (1988) Genes Dev. 2, 1570–1581.
[14] Kissinger, C.R., Liu, B., Martin-Blanko, E., Kornberg, T.B. and Pabo, C.A. (1991) Cell 63, 579–590.
[15] Wolberger, C., Vershon, A.K., Liu, B., Johnson, A.D. and Pabo, C.O. (1991) Cell 67, 517–528.
[16] Pomerantz, J.L. and Sharp, P.A. (1994) Biochemistry 33, 10851–10858.
[17] Klemm, J.D. and Pabo, C.O. (1996) Genes Dev. 10, 27–36.
[18] Verrijzer, C.P., Van Oosterhout, J.A.W.M., Van Weperen, W.W. and Van der Vliet, P.C. (1996) EMBO J. 10, 3007–3014.
[19] Ingraham, H.A., Flynn, S.E., Voss, J.W., Albert, V.R., Kapiloff, M.S., Wilson, L. and Rosenfeld, M.G. (1990) Cell 61, 1021–1033.
[20] Scott, M.P., Tamkun, J.W. and Hartzell, G.W. (1989) Biochim. Biophys. Acta 989, 25–48.