Review

Conflicting views on the molecular structure of the cancer antigen CA125/MUC16

Houda Bouanene[a,b,*] and Abdelhédi Miled[a,b]
[a]*Laboratory of Biochemistry and Molecular Biology, Faculty of Pharmacy, Monastir, Tunisia*
[b]*Laboratory of Biochemistry, CHU Farhat Hached Sousse, Tunisia*

Abstract. CA125 is a tumor antigen used to monitor the progression and regression of epithelial ovarian cancer. Despite the widespread use of CA125, the biochemical and molecular nature of this antigen is poorly understood. Analysis of the structure of CA125 is essential for determining the physiological role of this very significant tumor marker. Accumulated experimental evidence has shown that CA125 epitopes reside on a molecule of very complex architecture in terms of both protein backbone and oligosaccharide structures. It is not clear whether the heterogeneity of CA125 molecular characteristics are due to the variability of biological sources from which the molecule was isolated or to the different biophysical methods used for the characterization of all the oligosaccharides linked to CA125 or to the presence of glycoisoforms for this protein. This review attempts to summarize emerging data related to molecular characteristics of CA125 and to compare approaches undertaken to reach a better understanding of molecular features of this tumor marker.

Keywords: CA125 antigen, MUC16, glycosylation

1. Introduction

CA125 is a significant tumor marker associated with many human cancers but it is most widely used for diagnosis [1], monitoring of patients with ovarian cancers and for detection of progression of epithelial ovarian cancer [2]. This biomarker denotes the presence of ovarian and reproductive cancers in women, with serum concentrations of CA125 greater than 35 U/ml considered indicative of potential malignancies [3]. Interpretation of rising CA125 levels in diagnosed patients continues to be a topic of conversation among clinicians [4]. The CA125 is elevated in other cancers including endometrial [5], breast [6], lung [7], liver, stomach and bladder [8]. Despite the widespread use of CA125, the biochemical and molecular nature of this antigen is poorly understood. The knowledge of the domain structure of CA125 provides researchers with many possible explanations for the discrepancies in measurement of this tumor marker [9].

However, it's very high molecular weight and mucinous nature pose major obstacles for performing structural characterization studies. Here, we review emerging data related to molecular characteristics of CA125 and compare different approaches undertaken to reach a better knowledge about the structure of this tumor marker. Information on the expression, the molecular and biochemical nature of the CA125 antigen is of special interest since it could lead to improve assay methods as well as to a better understanding of its biological role in normal and malignant tissues.

2. What is known about CA125 as a clinical tumor marker?

The role of the cancer antigen CA125 in relation to ovarian cancer has been studied extensively. Char-

*Corresponding author: Houda Bouanene, Laboratory of Biochemistry and Molecular Biology, Faculty of Pharmacy- 5000, Moanstir, Tunisia. Tel.: +216 73 21 95 03; Fax: +216 73 22 67 02; E-mail: bouanenehouda2@yahoo.fr.

acterized in 1981 [10], the CA125 antigen remains the best-characterized serologic tumor marker usually adopted to evaluate the clinical situation in ovarian cancer patients. It has been used for many years to detect ovarian cancer recurrence and to monitor treatment response [11]. Large efforts have been deployed in search of biomarkers for early cancer detection, for disease classification, and for monitoring response to cancer therapy [12–15]. Various groups have investigated by combination approaches to increase the diagnostic value of multiple markers. Their sensitivities and specificities were often improved by combinations of two or more markers, usually with CA125 as one of these [16–18]. Approximately more than 30 serum markers have been evaluated alone or in combination with CA125 by different investigators [19]. Some of the most promising include HE4, mesothelin, M-CSF, OPN, KLK(s), and soluble EGF receptor [20]. Yet, no single marker has emerged as a serious competitor for CA125 [21]. The epithelial sialomucins are currently the most important group of markers for ovarian cancer [18]. These are heavily glycosylated proteins with relatively high molecular weights and include the well approved marker for monitoring ovarian cancer, CA125, also known as MUC16.

Clinical management of ovarian cancer patients is facilitated by CA125 determinations in serum. Presently, several assay systems based on different concepts and different methodologies are available to measure CA125. The first method developed to measure CA125 was a radioimmunometric assay that used the murine monoclonal antibody OC125 as both capture and indicator antibodies [22]. OC125 was obtained after immunization with the OVCA 433 cell line, which was derived from the ascites fluid of a patient with papillary cystadenocarcinoma of the ovary [10]. Monoclonal antibodies raised against other epitopes expressed by this molecule have led to the development of the CA125-II assay, a variant of the original CA125 assay that utilizes the M11 monoclonal antibody as the capture antibody [11,23]. Several studies document the similarities of sensitivity and specificity obtained with the original CA125 and newer CA125-II assays [24,25]. For the quantitative determination of the CA125 concentration in human serum, the microparticle enzyme immunoassay (MEIA) had a high degree of practicability and a short turn-around-time [26]. A fluorescent immunoassay was recently developed employing the ALYGN-SA antibody-orientation system that has been used to measure CA125 levels. This system displayed significantly increased sensitivity and should permit better identification and monitoring of ovarian cancer [3].

To investigate if the source of CA125 influenced epitope characterization, CA125 derived from ascites fluid and cervical mucus was assessed. Ascites-derived CA125 varied in size from about 190 to about 2.700 kD. Cervical mucus-derived CA125 treated with ultrasound changed its apparent size from more than 20.000 to 700 kD. Epitope mapping of antibodies was not grossly influenced by the size or source of CA125 used as target but different responses were depending on the antibody pair used [27].

3. What is known about CA125 in tissues?

CA125 was described previously as the prominent marker for ovarian cancer monitoring, however, the peptide epitope CA125 carried by MUC16 [gene mucin 16], is also present in normal physiological conditions [28,29]. MUC16 was expressed in normal tissues such as the surfaces of respiratory tract epithelia and female reproductive tract epithelia [30]. Of particular interest is the expression of MUC16 by the human ocular surface epithelia [31]. MUC16/CA125 was detected in the lacrimal gland [32], in the tear fluid [33], in the tips of epithelial processes of cornea and conjunctiva [31]. MUC16 was considered as a multifunctional mucin on the corneal epithelial surface. It has been suggested that MUC16 play a role in protecting the apical surface from pathogen adhesion [30]. Another postulated function of MUC16 is to alter the precorneal tear film and consequently it may be involved in dry eye pathology [34].

In ovarian cancer, MUC16/CA125 may facilitate tumor cells interactions. In this context, Galectin-1 [35] and mesothelin [36,37] have been reported to bind to CA125 from ovarian cancer cell lines. Owing to its nature as a glycoprotein, potential ligands of CA125 include lectins of the extracellular matrix (ECM) such as the β-galactoside-specific family of galectins [38]. CA125 cell surface expression might affect cell attachment to the extracellular matrix in a galectin-1-dependent manner [35]. Moreover, CA125 in tumor cells is implicated as a mediator of cellular adhesion, acting as a lubricant as well as a binder of mesothelin, a known cell-surface molecule that is expressed in the mesothelial lining of various body cavities. The binding of MUC16 to mesothelin may allow cellular adhesion activity, thus enabling tumors secreting CA125 to invade surrounding tissue and organs and consequently it helps the ovarian tumour cells to metastasize and form aggregates [36]. Thus, blocking the mesothelin-CA125 interaction may prevent or reverse metastasis and lead to overall improved survival in cancer patients [39].

Table 1
Differences in molecular sizes of CA125 isolated from different sources

Source of CA125		Methods to isolate CA125		Molecular weights
*Sera of ovarian carcinoma *Human milk		*Molecular size exclusion chromatography patients *Buoyant density ultracentrifugation		200 to 10^3 kDa
Ascites		*SDS-PAGE *HPLC gel filtration		*400 kDa *10^3 kDa
		*OC125 immunoaffinity chromatography *SDS-PAGE		
Ovarian cancer cell lines	OVCA433	*Molecular size exclusion chromatography *Buoyant density ultracentrifugation *Immunoblot (OC125)		200 to 10^3 kDa
		*OC125 immuno-affinity chromatography	SDS-PAGE under non denaturing and non reducing conditions	*10^3 kDa *200 to 600 kDa
			SDS-PAGE under denaturing and non reducing conditions	200 kDa
		*HPLC gel filtration		*400 kDa *10^3 kDa
		*OC125 immunoaffinity chromatography *SDS-PAGE		
	OVCAR-3	*VK-8 immuno-affinity chromatography *HPLC gel filtration *SDS-PAGE under reducing conditions		> 10^3 kDa
		*OC125 immuno-affinity chromatography *Electrophoretic analysis		200 kDa
	Caov-4 and NIH: OVCAR-3	*Gel filtration *Chromatography on wheat-germ-lectin-Sepharose *Western blot		> 200 kDa

4. Biochemical Analysis of CA125: No consensus about its molecular characteristics

Despite extensive and fundamental research, controversy still exists on the molecular characteristics of CA125. The biochemical determination of the CA125 antigen has produced conflicting data with evidence suggesting that it is a protein, a mucin, a carbohydrate, or a membrane-associated glycoprotein [40]. The first biochemical analysis of the CA125 antigen, which was purified from human milk, suggested that it was a sialylsaccharide antigen carried on a mucin glycoprotein [41]. Based on the gene structure, it is suggested that CA125 antigen is a transmembrane protein, comprising an extracellular domain, a transmembrane region, and an intracellular domain [42].

Experimental approaches for both the isolation and characterization of the antigen CA125 differ between studies. A comparison between different techniques and biological sources used to purify and characterize the CA125 and their results will lead to a greater understanding of the various data related to this tumor marker. The question is whether the biochemical differences found are resulting from the variability of biological sources from which the molecule was isolated (first hypothesis) or from the use of different techniques that can provide complete or incomplete characterization of all the oligosaccharides linked to CA125 (second hypothesis) or that this protein has different isoforms that are differentially expressed from one biological source to another or in the same source (third hypothesis)?

4.1. What is known about CA125 isolated from various biological fluids?

The fact that CA125 is a heavily glycosylated, high-molecular-weight protein has been established by a number of studies showing differences in molecular sizes of this tumor marker from different sources (Table 1). Davis et al. carried out an initial characterization of CA125 antigen isolated from a number of sources. The CA125 isolated from the sera of ovarian carcinoma patients was shown by gel electrophoresis, molecular size exclusion chromatography and buoyant density ultracentrifugation to have similar immunological and physical characteristics to the antigen isolat-

ed from human milk. A composite SDS gel resolved CA125 from the two sources of antigen into disperse bands of similar electrophoretic mobilities with apparent masses of 200 to 10^3 KDa to 1 million Daltons [43]. The results of immunoaffinity chromatography to isolate CA125 from ascites fluid of ovarian cancer patients indicate that CA125 moieties from patient ascites exist as high molecular weight glycoprotein complexes that migrate as 3 or 4 disperse bands with apparent molecular weight between 400 and 10^3 KDa and 1,000,000 Da using SDS-PAGE. It could not be determined whether these moieties represent oligomers of a smaller subunit, different glycoforms, or macromolecular complexes of distinct glycoproteins [44]. SDS-PAGE and gel filtration characterization showed that the purified antigen CA125 from human ascites exists as a high molecular weight complex, of up to 1.5 million daltons, which could be dissociated under strong denaturing conditions, giving rise to moieties with an apparent MW of 205 and 55 kD [45].

Analysis of CA125 from ascites, normal abdominal fluid, cyst fluid from a woman with a benign ovarian cyst, and serum from an ovarian cancer patient showed marked differences in the presence of these different components [40] and also as analyzed by SDS–PAGE and Western blotting of material from a variety of sources [46]. CA125 was found in normal abdominal fluid and cervical mucus with a molecular mass ranged between 1–4 MDa [40,47]. Endogenous protease activity and continuous deglycosylation of oligosaccharide chains in body fluids are thought to be responsible for the heterogeneity of CA125 in both size and charge [48]. Different sites of glycosylation on CA125 produce a heterogeneous mixture of varying molecular mass (mostly > 1000 kDa) glycoproteins. This heterogeneity shown by anionic exchange chromatography may also be due to differences in sialic acid content as a result of normal protein turnover and deglycosylation in body fluids [49] as well as to alternative splicing resulting in a variable number of tandem repeats [50]. Another possible reason for differences between CA125 molecules should be noted. Biological differences between cancer-associated and normal forms of the protein may exist. Further studies are underway to investigate the biological differences between these molecules. There is a need to implement strategies to share and to integrate data from analyses of various sources, including tumour-cell populations, tumour tissue, proximal fluids and peripheral blood, to gather all properties of MUC16/CA125. To achieve such a goal, this review gives a snapshot of various advances that have taken place in the field of MUC16/CA125 biomarker discovery.

4.2. Isolation of CA125 from different ovarian carcinoma cell lines

Many groups have focused on characterizing CA125 from various ovarian cancer cell lines. Here, we report the potential differences found in CA125 isolated from these ovarian cancer cell lines (Table 1).

By using a new murine monoclonal antibody (MAb VK-8) detecting the CA125, Lloyd et al. have purified the antigen from OVCAR-3 ovarian cancer cells by affinity chromatography. After CA125 purification based on VK-8 antibody, the biochemical characterization of this purified antigen showed that it had a high content of serine, threonine and proline [51]. This composition profile is very similar to other mucin-type molecules [52]. Treatment of the antigen with alkaline borohydride resulted in a loss of 47% of the serine content and 31% of the threonine, indicating that many of the serine and threonine residues were O-glycosylated with carbohydrate chains. This property is another characteristic of mucins [53]. The major conclusion from the mentioned study was that the biochemical properties of the purified antigen are characteristic of a mucin-type glycoprotein:

1. the molecule is highly glycosylated (77% w/w), mainly with galactose, N-acetylglucosamine, and N-acetylgalactosamine,
2. the protein moiety is rich in serine, threonine and proline,
3. many of the serine and threonine residues are glycosylated,
4. the glycan chains are almost entirely O-linked
5. these chains carry fucosylated Type 2 (Ley and Lex, and H type 2) blood group structures.

In aqueous buffer as well as in urea, the antigen exhibited a very high molecular weight ($> 10^3$ kDa), but it was degraded by proteolytic enzymes to smaller fragments that no longer reacted with the antibody [51]. The same cancer cell line (OVCAR-3) and the monoclonal antibody OC 125 were used in the study of Murdoch et al. for immunopurification of CA125 [54]. Electrophoretic analysis revealed a prevalent band of approximately 200 kDa.

NIH cancer cell line: OVCAR-3 and Caov-4 cells are other cancer cell lines that were the biological sources used to purify the antigen CA125 by gel filtration and chromatography on wheat-germ-lectin-Sepharose. CA125 purified by ammonium sulfate precipitation and gel filtration showed six bands in SDS gel. Using western blot, only one band above 200 kDa was stained [55].

The early form of CA125 isolated from NIH: OVCAR-3 ovarian cancer cell line was shown to be a glycoprotein, and no smaller species of mature forms of CA125 were formed, suggesting that CA125 was synthesized through a glycosylated 400-kDa precursor [56]. Concerning the ovarian cancer cell line SHIN-3, a molecule of 49 kDa was secreted [57].

Analysis of 35S-methionine-labeled CA125 produced in the amnion derived WISH cell line showed similar synthesis of CA125 within two hours as a significantly larger than 200-kDa product, which was secreted into the medium within 4 h as a 200-kDa protein. This secreted form was thought to be an underestimate of the true size of the newly synthesized protein and it is a result of proteolytic degradation upon secretion [49].

Using SDS-PAGE, CA125 isolated from human ovarian carcinoma cell line OVCA433 by HPLC gel filtration column and then by OC125 immunoaffinity chromatography showed 3 or 4 disperse bands with apparent molecular weight between 400 and 10^3 kDa [44]. In another study, isolation of CA125 antigen of higher purity from OVCA433 supernatant was achieved by a series of steps including OC125 immunoaffinity chromatography. Under non denaturing and non reducing conditions, CA125 shows high molecular mass of greater than 1 million Da and lower molecular mass of 200.000 to 600.000 Da. Under denaturing and non reducing conditions CA125 showed Molecular mass of 200.000 Da. It was concluded that CA125 epitopes were carried on large glycoproteins, with molecular weight of 2×10^5–10^6 Da, which did not have the properties of a typical mucin and that CA125 is an extremely complex molecule from both the proteomic and glycomic perspectives [43].

5. After more than 20-years effort to characterize the antigen CA125

Considerable efforts have been deployed towards characterizing the antigen CA125. Progress in understanding the molecular nature of CA125 has been hampered by difficulties encountered in attempting to clone this antigen. Direct amino acid sequencing of CA125 is complicated by the large size and heavy glycosylation of the protein. Expression cloning provides a convenient alternative to determine the primary sequence of CA125. The first cDNA clones for CA125 were published the same year by two independent groups [42,58]. Using the same cancer cell lines and protocols to purify CA125 as in 1997, Lloyd et al. have cloned, by expression cloning, a long partial cDNA sequence corresponding to a new mucin species (designated MUC16).

MUC16 belongs to a family of membrane-bound mucins that encodes a transmembrane-bound molecule called CA125. Mucin glycoproteins are a heterogeneous family of high-molecular-mass, heavily glycosylated proteins differentially expressed in epithelial tissue. To date, nineteen human apomucin genes have been identified (*MUC1-MUC19*) and the corresponding fully glycosylated mucins (up to 80% carbohydrate) are designed according to them [59]. However, each mucin can be distinguished by the length and amino acid sequence of its tandem repeat unit [60].

Confirmation that the CA125 clone represents a mucin is provided by the presence of tandem repeat units rich in threonine and serine [58]. A domain of tandem repeat arrays enriched in hydroxyl amino acids characterizes all known mucins. The abundance of threonine and serine residues provides many potential sites for O-glycosylation [61]. MUC16 has two particularities from other mucins: the high content of leucine and the unusual length of the repeat units (156 amino acids) that characterizes the extracellular region. This extracellular region is dominated by residues that carry multiple $O-$ and $N-$glycans. Amino acid analysis of MUC16 revealed that its sequence contains numerous lysine and arginine residues that are remote from the postulated O-glycosylation sites [58]. The crystal structure of the repeats was characterized, revealing the presence of varying types of SEA (S = sea urchin, E = enterokinase, and A = agrin) domains, a structural motif found in highly O-glycosylated membrane proteins that is cleaved and then reassociates via noncovalent bonds [62].

The tandem repeat sequence is recognized by the antibodies OC125 and M11 confirming that the detection mechanism of most commercial assays for CA125 is found in the extracellular repeat domains [58]. Variation in the MUC16 gene may cause changes in the CA125 protein. Despite improvements in the specificity of the CA125 immunoassay, in most circumstances, serum laboratory tests detect one molecule and quantify it, with no genotypic information required [9].

It is now recognized that the CA125 molecule is a truly massive transmembrane glycoprotein consisting of multiple (60+) repeat domains and an amino terminus. The extracellular portion is highly glycosylated (Fig. 1) [42]. Northern blot analysis indicates that the mRNA (GeneBank accession number AF414442.2) is approximately 66765 bp. This vast message encodes
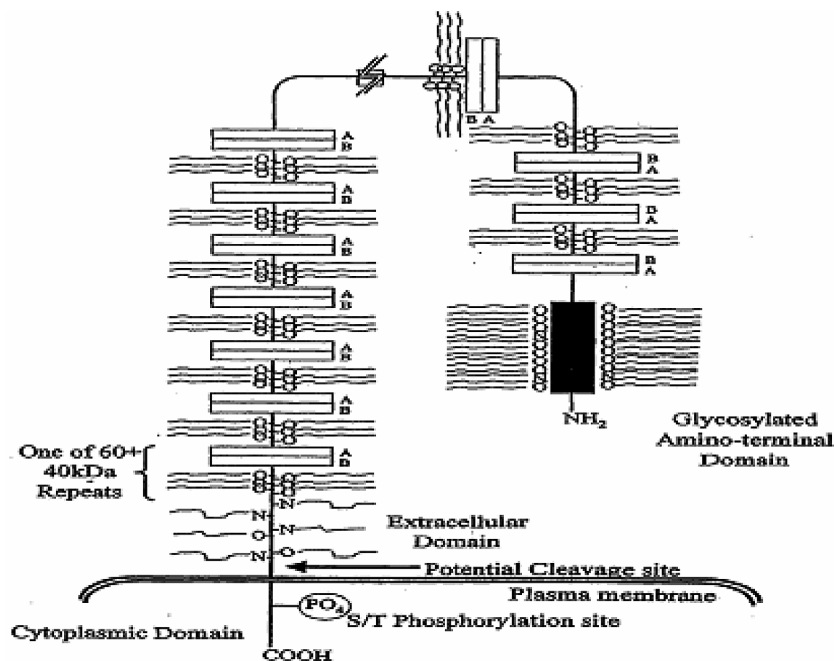
Fig. 1. Proposed structure of CA125 by *O'Brien et al. 2001*.

a theoretical product of 22152 amino acids [63]. Due to the highly iterative nature of the repeat region both nucleotide and translated amino acid sequences contain significant numbers of unknown residues. The CA125 mRNA was first characterized as a 5.8-kb transcript before 5' additions that extended the sequence to its present form [42,58,63]. The CA125 molecule is composed of an N-terminal domain, a tandem repeat region and a cytoplasmic domain and there is a typically high content of proline, serine and threonine [58]. The cytoplasmic tail of MUC16 is not well characterized; it is relatively short, consisting of 256 amino acids [42]. It contains several potential tyrosine phosphorylation sites, which are conserved in the mouse homolog of MUC16. Although CA125 is conserved in some mammals [64,65] its study is hampered by its lack of conservation in rodents. The absence of CA125 in rodents may be due to ovarian biological differences. Thus it is difficult to extrapolate data concerning CA125 structure from mouse to human [66].

The heterogeneity of the protein backbone and glycosylation distribution make difficult to estimate an average molecular weight of CA125/MUC16. The core protein sequence of this protein could have a mass approaching 2.5 million Da and the very high degree of glycosylation suggests that the average molecular mass of CA125 may be 3.5 million Da. Although the initial data indicated that the N-terminal domain is made up

of 1637 amino acids, recently, it has been shown to be vastly larger, i.e. an additional 10431 amino acids have been found. The tandem repeat domain consists of 40–60 repeats, each 156 amino acids long with a disulphide bridge loop of 19 amino acids [42–57,59, 59–63]. Although the overall structure is well conserved, only few repeats have identical amino acid sequences. Of additional note, MUC16 codes splicing variants that give rise to polypeptides that differ in the number of SEA domains. Consequently, CA125 antigen may have 7, 20, or 60 tandem repeats (TR) increasing the core molecular weight from 200000 daltons to almost 4 million daltons [62].

Besides its putative nature as an integral membrane protein, soluble fragments of CA125 have been observed [58,67]. The N-terminal subunit of MUC16 is a typical heavily O glycosylated mucin domain. Numerous potential N glycosylation sites (Asn-X-Ser/Thr, where X is any amino acid except Pro) were also found in the sequence, including two that are perfectly conserved in the TR region within the mucin [58].The C-terminal portion consists of a 32-residue CT with several tyrosine, serine, and threonine sites for potential phosphorylation [42,67].

A few years later, Wong et al. [49] provided an unambiguous sequencing of CA125-associated glycans that required the use of rigorous biophysical methods. In their detailed analysis, they have circumvent-

ed many problems associated with biochemical analysis of CA125. The data also confirmed robust N-glycosylation of CA125, supporting previous studies suggesting this possibility [48,69]. The Wong group criticized previous approaches used by Lloyd et al. who employed indirect approaches for characterizing the oligosaccharides associated with CA125 derived from OVCAR-3 cells [51]. Because of the relatively insensitive methods of detection that were employed, only a limited number of oligosaccharides were analyzed and they have identified only several O-linked glycans expressed on CA125. The obtained results did not provide complete characterization of all the oligosaccharides linked to CA125. Thus, experimental design has a crucial role in a successful search leading to the highly confident identification of MUC16/CA125.

In a latest study, CA125 antigen was isolated from first trimester placental extract by gel filtration, ion-exchange chromatography and lectin affinity chromatography. Elution profiles from the lectin-affinity columns demonstrated molecular subpopulations bound with low, intermediate and high affinity. Thus, CA125 antigen was found to be heterogeneous in respect to molecular mass species and to the existence of different glycoisoforms corresponding to molecular subpopulations. These data was obtained using proteomic analysis and glycobiochemical characterization of the isolated antigen [70].

More recently, a comparison of the glycosylation of CA125 antigen from amniotic fluid and ovarian carcinoma cell line OVCAR-3 was done in order to detect possible differences as a specific marker of their origin. The findings from this investigation were particularly interesting suggesting that the degree and nature of glycosylation depend on the origin of the glycoprotein. There was also a marked difference in carbohydrate composition between CA125 antigens isolated from each source suggesting the existence of multiple glycoforms, with both N- and O-linked glycan chains [71]. Investigations of glycoconjugates present on MUC16/CA125 protein are particularly interesting as they may provide information about the potential biological function of CA125 [36,37].

Like MUC1, a proadhesive function has been attributed to MUC16 owing to its ability to bind specifically to human galectin-1 with high efficiency [35] and to the tumor marker mesothelin [36]. The interaction between CA125 and mesothelin is capable of enhancing ovarian cancer cell metastasis. The binding of MUC16/CA125 to the N-terminus of the extracellular domain of mesothelin, that is sufficient and necessary for binding to CA125, requires N-linked oligosaccharides of this mucin. It has been found that the extraordinarily abundant N-glycans on CA125, presumably in the tandem repeat region, are required for binding to both glycosylated and non-glycosylated mesothelin [37]. In fact, the mesothelin binding site on CA125 may lie within the 156 amino acid tandem repeat units, indicating multimeric binding of mesothelin to CA125 [39]. The importance of MUC16 glycosylation was also shown in an investigation of the interaction of CA125-CTERM with galectin-1. These data establish that the interaction of CA125-CTERM with galectin-1 largely depends on O-linked β-galactose-terminated oligosaccharide chains. Moreover, the cellular context in which CA125 is expressed has a significant impact on its binding specificity for members of the galectin family [35].

Despite intensive research efforts, molecular structure of MUC16/CA125 remains controversial. Thus, the challenge of discovering MUC16/CA125 structure necessitates integrated and cooperative efforts. Recent studies have been undertaken by investigators who have focused on analysis of this biomarker from different biological sources by proteomic approaches using the MALDI-TOF mass spectrometry [70,72], the tandem mass spectrometry [MS/MS] [68,73] or the nano-LC-FT-ICR MS [74,75]. Although mucins are difficult to identify with proteomics as most of their peptides are glycosylated and thus not identified in the protein databases [72], implementation of such approaches is likely to yield more information about the complexity of MUC16/CA125 and to achieve a substantial depth of analysis. Given the heterogeneity in the approaches applied and the sampling nature of MS analysis, partial identification of various peptides of MUC16/CA125 was the major result. Across these studies, most of the peptides identified, occur several times in the MUC16 sequence and correspond to tandem repeat peptides [35, 72–75]. Efforts done to reduce discrepancy related to CA125/MUC16 structure may be noticeably improved with the development of proteomic technologies.

6. Conclusion

The present observations highlight the need for careful selection of approaches used to identify such a huge glycoprotein in term of peptide core and glycosylation. Thus, reviewing the biological and technological challenges embodied for the identification of MUC16/CA125 may provide an opportunity to know

the successes and the failures of existing biological foundation of such biomarker. Moreover, the relatively significant progress that has been made in the last 20 years highlights a conclusive evidence for the heterogeneity of the MUC16/CA125 molecule.

In conclusion, it is obvious that CA125 can exist in a number of molecular mass forms. In the light of the different results previously mentioned, three hypotheses [biochemical differences related to CA125 structure may result from the variability of biological sources from which the molecule was isolated or from the use of different techniques that can provide complete or incomplete characterization of all the oligosaccharides linked to CA125 or that this protein has different isoforms that are differentially expressed from one biological source to another or in the same source], suggested in this review are likely to be true. The debate will no doubt continue in order to gain a better knowledge about the molecular features of CA125. This is a crucial step in leading to a better understanding of its function in normal and pathological conditions.

Conflict of interest

We declare that we have no conflict of interest.

References

[1] T. Maggino and A. Gadducci, Serum markers as prognostic factors in epithelial ovarian cancer: an overview, *Eur J Gynaecol Oncol* **21** (2000), 64–69.

[2] G.J. Rustin, M. Marples, A.E. Nelstrop, M. Mahmoudi and T. Meyer, Use of CA-125 to define progression of ovarian cancer in patients with persistently elevated levels, *J Clin Oncol* **19** (2001), 4054–4057.

[3] D. Sok, L.J.A. Clarizia, L.R. Farris and M.J. McDonald, Novel fluoroimmunoassay for ovarian cancer biomarker CA-125, *Analytical and bioanalytical chemistry* **393** (2009) 1521–1523.

[4] M.K. Tuxen, G. Soletormos, G.J.S. Rustin, A.E. Nelstrop and P. Dombernowsky, Biological variation and analytical imprecision of CA125 in patients with ovarian cancer, *Scand J Clin Lab Investigation* **60** (2000), 713–722.

[5] B. Yu, P.Z. Xu, Q.W. Wang, H. Zhou and H.X. Zhou, Clinical value of tumour specific growth factor (TSGF) and carbohydrate antigen-125 (CA-125) in carcinoma of the endometrium, *J Int Med Res* **37** (2009), 878–883.

[6] L.F. Norum, B. Erikstein and K. Nustad, Elevated CA125 in breast cancer – A sign of advanced disease, *Tumor Biol* **22** (2001), 223–228.

[7] S. Homma, H. Satoh, K. Kagohashi, M. Fujiwara, H. Kamma and K. Sekizawa, Production of CA125 by human lung cancer cell lines, *J Int Med Res* **37** (2009), 878–883.

[8] C. Miralles, M. Orea, P. Espana, M. Provencio, A. Sánchez, B. Cantos, R. Cubedo, E. Carcereny, F. Bonilla and T. Gea, Cancer antigen 125 associated with multiple benign and malignant pathologies, *Ann Surg Oncol* **10** (2003), 150–154.

[9] M.R. McLemore and B. Aouizerat, Introducing the MUC16 Gene: Implications for Prevention and Early Detection in Epithelial Ovarian Cancer, *Biological Res Nurs* **6** (2005), 262–267.

[10] Jr.R.C. Bast, M. Feeney, H. Lazarus, L.M. Nadler, R.B. Colvin and R.C. Knapp, Reactivity of a monoclonal antibody with human ovarian carcinoma, *J Clin Investig* **68** (1981), 1331–1337.

[11] Jr.R.C. Bast, F.J. Xu, Y.H. Yu, S. Barnhill, Z. Zhang and G.B. Mills, CA 125: the past and the future, *Int J Biol Markers* **13** (1998), 179–187.

[12] D.G. Rosen, L. Wang, J.N. Atkinson, Y. Yu, K.H. Lu, E.P. Diamandis, I. Hellstrom, S.C. Mok, J. Liu and R.C. Bast, Potential markers that complement expression of CA125 in epithelial ovarian cancer, *Gynecol Oncol* **99** (2005), 267–277.

[13] N. Scholler, M. Crawford, A. Sato, C.W. Drescher, K.C. O'Briant, N. Kiviat, G.L. Anderson and N. Urban, Bead-based ELISA for validation of ovarian cancer early detection markers, *Clin Cancer Res* **12** (2006), 2117–2124.

[14] R.G. Moore, A.K. Brown, M.C. Miller, S. Skates, W.J. Allard, T. Verch M. Steinhoff, G. Messerlian, P. DiSilvestro, C.O. Granai and R.C. Jr. Bast, The use of multiple novel tumor biomarkers for the detection of ovarian carcinoma in patients with a pelvic mass, *Gynecol Oncol* **108** (2008), 402–408.

[15] Y. Zheng, D. Katsaros, S.J. Shan, I.R. De la Longrais, M. Porpiglia, A. Scorilas, N.W. Kim, R.L. Wolfert, I. Simon, L. Li, Z. Feng and E.P. Diamandis, A multiparametric panel for ovarian cancer diagnosis, prognosis, and response to chemotherapy, *Clin Cancer Res* **13** (2007), 6984–6992.

[16] C. Crump, M.W. McIntosh, N. Urban, G. Anderson and B.Y. Karlan, Ovarian cancer tumor marker behavior in asymptomatic healthy women: Implications for screening, *Cancer Epidemiol Biomarkers Prev* **9** (2000), 1107–1111.

[17] Z. Zhang, Jr.R.C. Bast, Y. Yu, J. Li, L.J. Sokoll and A.J. Rai, Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer, *Cancer Res* **64** (2004), 5882–5890.

[18] K.L. Terry, P.M. Sluss, S.J. Skates, S.C. Mok, B. Ye, A.F. Vitonis and D.W. Cramer, Blood and urine markers for ovarian cancer: A comprehensive review, *Dis Markers* **20** (2004), 53–70.

[19] R.C. Jr. Bast, D. Badgwell, Z. Lu, R. Marquez, D. Rosen, J. Liu, K.A. Baggerly, E.N. Atkinson, S. Skates, Z. Zhang, A. Lokshin, U. Menon, I. Jacobs and K. Lu, New tumor markers: CA125 and beyond, *Int J Gynecol Cancer* **15** (2005), 274–281.

[20] N. Kothandaraman, C. Zhao and C. Mahesh, Ovarian cancer proteomics: Many technologies one goal, *Proteomics Clin Appl* **2** (2008), 195–218.

[21] A.J. Rai, A.J. Zhang, Z. Rosenzweig, J. Shih, L. Pham, T. Fung, J. Lori, L.J. Sokoll and D.W. Chan, Proteomic Approaches to Tumor Marker Discovery Identification of Biomarkers for Ovarian Cancer, *Arch Pathol Lab Med* **126** (2002), 1518–1526.

[22] R.C. Jr. Bast, T.L. Klug, E. St John, E. Jenison, J.M. Niloff, H. Lazarus, R.S. Berkowitz, T. Leavitt, C.T. Griffiths, L. Parker, V.R. Zurawski and Jr.R.C. Knapp, A radioimmunoassay using a monoclonal antibody to monitor the course of epithelial ovarian cancer, *N Engl J Med* **309** (1983), 883–887.

[23] K. Nustad, R.C. Jr. Bast, T.J. Brien, O. Nilsson, P. Seguin, M.R. Suresh, T. Saga, S. Nozawa, O.P. Bormer, H.W. de Brui-

jn, M. Nap, A. Vitali, M. Gadnell, J. Clark, K. Shigemasa, B. Karlsson, F.T. Kreutz, D. Jette, H. Sakahara, K. Endo, E. Paus, D. Warren, S. Hammarstrom, P. Kenemans and J. Hilgers, Specificity and affinity of 26 monoclonal antibodies against the CA125 antigen: first report from the ISOBM TD-1 workshop International Society for Oncodevelopmental Biology and Medicine, *Tumor Biol* **17** (1996), 196–219.

[24] K. Tamakoshi, F. Kikkawa, N. Hasegawa, H. Ishikawa, K. Mizuno, M. Kawai and Y. Tomoda, Clinical value of a new serum tumor marker, CA125II, in gynecologic disease: comparison with CA125, *Gynecol Obstet Investig* **39** (1995), 125–129.

[25] E.M. Davelaar, G.J. Van Kamp, R.A. Verstraeten and P. Kenemans, Comparison of seven immunoassays for the quantification of CA 125 antigen in serum, *Clin Chem* **44** (1998), 1417–1422.

[26] O. Mogensen and B. Mogensen, Measurement of CA 125 by an automated microparticle enzyme immunoassay; comparison with a manual enzyme immunoassay, *Scan J Clin Lab Invest* **52** (1992), 507–512.

[27] K. Nustad, Y. Lebedin, K.O. Lloyd, K. Shigemasa, H.W.A. de Bruijn, B. Jansson, O. Nilsson, K.H. Olsen and T.J. O'Brien, Epitopes on CA125 from Cervical Mucus and Ascites Fluid and Characterization of Six New Antibodies, *Tumor Biol* **23** (2002), 303–314.

[28] F. Paulsen, Cell and molecular biology of human lacrimal gland and nasolacrimal duct mucins, *Int Rev Cytol* **249** (2006), 229–279.

[29] F. Paulsen and M. Berry, Mucins and TFF peptides of the tear film and lacrimal apparatus, *Prog Histochem Cytochem* **41** (2006), 1–56.

[30] T. Blalock, S. Spurr-Michaud, A. Tisdale, S. Heimer, M. Gilmore, V. Ramesh and I.K. Gipson, Functions of MUC16 in corneal epithelial cells, *Invest Ophthalmol Vis Sci* **48** (2007), 4509–4518.

[31] P. Argeso, S. Spurr-Michaud, C.L. Russo, A. Tisdale and I.K. Gipson, MUC16 mucin is expressed by the human ocular surface epithelia and carries the H185 carbohydrate epitope, *Invest Ophthalmol Vis Sci* **44** (2003), 2487–2495.

[32] K. Jäger, G. Wu, S. Sel, F. Garreis, L. Bräuer and F. Paulsen, Detection and distribution of MUC16 in the lacrimal apparatus, *Histochem Cell Biol* **127** (2006), 433–438.

[33] S. Spurr-Michaud, P. Argeso and I.K. Gipson, Assay of mucins in human tear fluid, *Exp Eye Res* **84** (2007), 939–950.

[34] F. Paulsen, K. Jäger, D. Worlitzschb, L. Bräuera, U. Schulzea and G. Schäferc, Regulation of, MUC16 by inflammatory mediators in ocular surface epithelial cell lines, *Ann Anatomy* **190** (2008), 59–70.

[35] C. Seelenmeyer, S. Wegehingel, J. Lechner and W. Nickel, The cancer antigen CA125 represents a novel counter receptor for galectin-1, *J Cell Sc* **116** (2003), 1305–1318.

[36] A. Rump, Y. Morikawa, M. Tanaka, S. Minami, N. Umesaki, M. Takeuchi and A. Miyajima, Binding of ovarian cancer antigen CA125/MUC16 to mesothelin mediates cell adhesion, *J Biol Chem* **279** (2004), 9190–9198.

[37] J.A.Gubbels, J. Belisle, M. Onda, C. Rancourt, M. Migneault, M. Ho, T.K. Bera, J. Connor, B.K. Sathyanarayana, B. Lee, I. Pastan and M.S. Patankar, Mesothelin-MUC16 binding is a high affinity, N-glycan dependent interaction that facilitates peritoneal metastasis of ovarian tumors, *Mol Cancer* **5** (2006), 50.

[38] G.A. Rabinovich, N. Rubinstein and L. Fainboim, Unlock the secrets of galectins: a challenge at the frontier of the glyco-immunology, *J Leukocyte Biol* **71** (2002), 741–752.

[39] O. Kaneko, L. Gong, J. Zhang, J.K. Hansen, R. Hassan, B. Lee and M. Ho, A Binding Domain on Mesothelin for CA125/MUC16, *J Biol Chem* **284** (2009), 3739–3749.

[40] K. Nustad, M. Onsrud, B. Jansson and D. Warren, CA125-epitopes and molecular size, *Int J Biol Markers* **134** (1998), 196–199.

[41] F.G. Hanisch, G. Uhlenbruck, C. Dienst, M. Stottrop and E. Hippauf, CA125 and CA 19-9: two cancer-associated sialylsaccharide antigens on a mucus glycoprotein from human milk, *Eur J Biochem* **149** (1985), 323–330.

[42] T.J. O'Brien, J.B. Beard, L.J. Underwood, R.A. Dennis, A.D. Santin and L. York, The CA125 gene: an extracellular superstructure dominated by repeat sequences, *Tumor Biol* **22** (2001), 348–366.

[43] H.M. Davis, V.R. Jr. Zurawski, R.C. Jr. Bast and T.L. Klug, Characterization of the CA125 antigen associated with human epithelial ovarian carcinoma, *Cancer Res* **46** (1986), 6143–6148.

[44] Y.H. Yu, D.M. Schlossman, C.L. Harrison, A. Rhinehardt-Clark, J.T. Soper, T.L. Klug, V.R. Zurawski and R.C. Jr. Bast, Coexpression of Different Antigenic Markers on Moieties That Bear CA 125 Determinants, *Cancer Res* **51** (1991), 468–475.

[45] M.T. De los Frailes, S. Stark, W. Jaeger, A. Hoerauf and L. Wildt, Purification and characterization of the CA 125 tumor-associated antigen from human ascites, *Tumor Biol* **14** (1993), 18–29.

[46] A. Barbati, Immunoblotting characterization of CA 125 in biological fluids: difference between pregnancy and cancer CA 125 origin, *Anticancer Res* **166B** (1996), 3621–3624.

[47] H.W. De Bruijn, C. Van-Beeck-Calkoen, S. Jager, J.M. Duk, J.G. Aalders and G.J. Fleuren, The tumor marker CA125 is a common constituent of normal cervical mucus, *Am J Obstet Gynecol* **154** (1986), 1088–1091.

[48] B. Macao, D.G.A. Johansson, G.C. Hansson and T. Härd, Auto-proteolysis coupled to protein folding in the SEA domain of the membrane-bound MUC1 mucin, *Nat Struct Mol Biol* **13** (2006), 71–76.

[49] T.J. O'Brien, H. Tanimoto, I. Konishi and M. Gee, More than 15 years of CA125: what is known about the antigen, its structure and its function, *Int J Biol Markers* **13** (1998), 188–195.

[50] M.R. McLemore and B. Aouizerat, Introducing the MUC16 Gene: Implications for Prevention and Early Detection in Epithelial Ovarian Cancer, *Biol Res Nur* **6** (2005), 262–267.

[51] K.O. Lloyd, B.W.T. Yin and V. Kudryashov, Isolation and characterization of ovarian cancer antigen CA 125 using a new monoclonal antibody VK-8: identification as a mucin – type molecule, *Int J Cancer* **71** (1997), 842–850.

[52] A. Gottschalk and A.S. Bhargava, Submaxillary gland glycoproteins. In: *Glycoproteins*, A. Gottschalk, ed., Part B, 1972, 810–829.

[53] K.O. Lloyd, Glycoproteins with blood group activity, In: *International Review of Science*, G.O. Aspinall, ed., *Organic Chemistry* **7** (1976), 252–281.

[54] W.J. Murdoch, E.A. Van Kirk and A.Mc. Smedts, Complement-inhibiting effect of ovarian cancer antigen CA-125, *Cancer Lett* **236** (2006), 54–57.

[55] B.C. Schultes, R.P. Baum, A. Niesen, A.A. Noujaim and R. Madiyalakan, Anti-idiotype induction therapy: anti-CA125 antibodies (Ab3) mediated tumor killing in patients treated with ovarex mAb B43.13 (Ab1) *Cancer Immunol Immunother* **46** (1998), 201–212.

[56] K.O. Lloyd and B.W.T. Yin, Synthesis and secretion of the ov-

arian cancer antigen CA125 by the human cancer cell line NIH: OVCAR-3, *Tumor Biol* **222** (2001), 77–82.

[57] S. Imai, H. Maeda, Y. Kiyozuka, T. Noda, I. Moriyama and M. Ichijo, Characterization of the CA125 antigen secreted from a newly established human ovarian cancer cell line (SHIN-3), *Acta Pathol Jpn* **39** (1989), 43–49.

[58] B.W.T. Yin and K.O. Lloyd, Molecular cloning of the CA125 ovarian cancer antigen, *J Biol Chem* **276** (2001), 27371–27375.

[59] E. Levi, D.S. Klimstra, N.V. Adsay, A. Andea and O. Basturk, MUC1 and MUC2 in pancreatic neoplasia, *J Clin Pathol* **57** (2004), 456–462.

[60] L.E. Vinall, S. Hilla, P. Pigny, W.S. Pratt, N. Toribara, J.R. Gum, Y.S. Kim, N. Porchet, J.P. Aubert and D.M. Swallow, Variable number tandem repeat polymorphism of the mucin genes located in the complex on 11p15.5, *Hum Genet* **102** (1998), 357–366.

[61] J.L. Desseyn, D. Tetaert and V. Gouyer, Architecture of the large membrane-bound mucins, *Gene* **410** (2008), 215–222.

[62] T. Maeda, M. Inoue, S. Koshiba, T. Yabuki, M. Aoki and E. Nunokawa, Solution structure of the SEA domain from the murine homologue of ovarian cancer antigen CA125 (MUC16), *J Biol Chem* **279** (2003), 13174–13182.

[63] T.J. O'Brien, J.B. Beard and L.J. Underwood, The CA 125 gene: A newly discovered extension of the glycosylated N-terminal domain doubles the size of this extracellular super-structure, *Tumor Biol* **23** (2002), 154–169.

[64] E.J. Nouwen, S. Dauwe and M.E. De Broe, Occurrence of the mucinous differentiation antigen CA125 in genital tract and conductive airway epithelia of diverse mammalian species (rabbit, dog, monkey), *Differentiation* **45** (1990), 192–198.

[65] A.C. McDonnel, E.A. Van Kirk, K.J. Austin, T.R. Hansen, E.L. Belden and W.J. Murdoch, Expression of CA-125 by progestational bovine endometrium: prospective regulation and function, *Reproduction* **126** (2003), 615–620.

[66] N. Scholler, B. Garvik, M. Hayden-Ledbetter, T. Kline and N. Urban, Development of a CA125-mesothelin cell adhesion assay as a screening tool for biologics discovery, *Cancer Lett* **247** (2007), 130–136.

[67] J.L. Fendrick, I. Konishi, S.M. Geary, T.H. Parmley, J.G. Jr. Quirk and T.J. O'Brien, CA125 phosphorylation is associated with its secretion from the WISH human amnion cell line, *Tumor Biol* **18** (1997), 278–289.

[68] N.K. Wong, R.L. Easton, M. Panico, M. Sutton-Smit, J.C. Morrison, F.A. Lattanzio, H.R. Morris, G.F. Clark, A. Dell and M.S. Patankar, Characterization of the Oligosaccharides Associated with the Human Ovarian Tumor Marker CA125, *J Biol Chem* **278** (2003), 28619–28634.

[69] A. Nagata, N. Hirota, T. Sakai, M. Fujimoto and T. Komoda, Molecular nature and possible presence of a membranous glycan-phosphatidylinositol anchor of CA125 antigen, *Tumor Biol* **12** (1991), 279–286.

[70] M.M. Jankovic and B.S. Tapuskovic, Molecular forms and mi-croheterogeneity of the oligosaccharide chains of pregnancy-associated CA125 antigen, *Hum Reprod* **20** (2005), 2632–2638.

[71] M.M. Jankovic and B.S. Milutinovic, Glycoforms of CA125 antigen as a possible cancer marker, *Cancer Biomark* **4** (2008), 35–42.

[72] B. Milutinovic and M. Jankovic, Analysis of the protein and glycan parts of CA125 antigen from human amniotic fluid, *Arch Biol Sci* **59** (2007), 97–103.

[73] J.R. Davies, S. Kirkham, N. Svitacheva, D.J. Thornton, I. Carl-stedt, MUC16 is produced in tracheal surface epithelium and submucosal glands and is present in secretions from normal human airway and cultured bronchial epithelial cells, *Int J Biochem Cell B* **39** (2007), 1943–1954.

[74] Y. Andersch-Bjorkman, K.A. Thomsson, J.M. Holme'n Lars-son, E. Ekerhovd and G.C. Hansson, Large Scale Identifica-tion of Proteins, Mucins, and Their O-Glycosylation in the Endocervical Mucus during the Menstrual Cycle, *Mol Cell Proteomics* **6** (2007), 708–716.

[75] D.J. States, G.S. Omenn, T.W. Blackwell, D. Fermin, J. Eng, D.W. Speicher and S.M. Hanash, Challenges in derinving high-confidence protein identification from data gathered by a HUPO plasma proteome collaborative study, *Nat Biotechnol* **24** (2006), 333–338.