

PUBLISHED VERSION

Weiland, Florian; Martin, Karina; Oehler, Martin Klaus; Hoffmann, Peter
[Deciphering the molecular nature of ovarian cancer biomarker CA125](#), International Journal of Molecular Sciences, 2012; 13(8):10568-10582.

© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).

PERMISSIONS

<http://www.mdpi.com/about/openaccess>

All articles published by MDPI are made available under an open access license worldwide immediately. This means:

- everyone has free and unlimited access to the full-text of all articles published in MDPI journals, and
- everyone is free to re-use the published material given proper accreditation/citation of the original publication.
- open access publication is supported by authors' institutes or research funding agency by payment of a comparatively low [Article Processing Charge \(APC\)](#) for accepted articles.

15th January 2013

<http://hdl.handle.net/2440/73821>

Review

Deciphering the Molecular Nature of Ovarian Cancer Biomarker CA125

Florian Weiland^{1,†}, **Karina Martin**^{1,†}, **Martin K. Oehler**² and **Peter Hoffmann**^{1,*}

¹ Adelaide Proteomics Centre, School of Molecular and Biomedical Science, University of Adelaide, Adelaide, SA 5005, Australia; E-Mails: florian.weiland@adelaide.edu.au (F.W.); karina.martin@adelaide.edu.au (K.M.)

² Research Centre for Reproductive Health, Robinson Institute, University of Adelaide, Adelaide, SA 5005, Australia; E-Mail: martin.oehler@adelaide.edu.au

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: peter.hoffmann@adelaide.edu.au; Tel.: +61-8-8313-5507; Fax: +61-8-8313-4362.

Received: 2 July 2012; in revised form: 3 July 2012 / Accepted: 13 August 2012 /

Published: 22 August 2012

Abstract: The ovarian cancer biomarker CA125 has been extensively investigated over the last 30 years. The knowledge about the exact molecular nature of this protein, however, remains fragmented. This review provides an overview of the structural research regarding CA125, and presents an orthogonal verification method to confirm the identity of this molecule. The need for independent identification of CA125 is exemplified by several reports where mutually exclusive data concerning the existence of isoforms and the glycan moieties is presented. Mass spectrometry can overcome the pitfalls of a single detection/identification method such as antibody probing. Independent verification of CA125 identity in characterization studies will help establish a refined model of its molecular structure that will promote the development of new approaches for diagnosis, prognosis and therapy of ovarian cancer.

Keywords: CA125; MUC16; ovarian cancer; biomarker; mass spectrometry

1. Introduction

The discovery of CA125 represented a milestone for ovarian cancer detection. In 1981, Bast *et al.* identified a murine monoclonal antibody, OC125, which reacted almost exclusively with ovarian cancer cell lines and cryo-preserved tissue of ovarian cancer patients [1]. An immuno assay for the detection of the OC125 antigen (CA125) in serum was soon developed and demonstrated a significant correlation between CA125 expression levels and the regression, stability or progression of epithelial ovarian carcinomas [2]. However, elevated serum concentrations were also found in 29% of patients with non-gynecological cancers as well as individuals presenting with benign conditions such as endometriosis, menstruation and pregnancy [2,3]. Due to the lack of specificity of CA125 a risk of malignancy index (RMI) is commonly used in clinical practice which combines CA125 levels, ultrasonographic findings and menopausal status of the patient. Using this approach 85% sensitivity and 97% specificity can be achieved when distinguishing benign from malignant ovarian disease [4]. However, this is not a suitable screening strategy for ovarian cancer as this disease has a low prevalence and requires a specificity of more than 99.6% to achieve an acceptable positive predictive value of 10% [5]. A large randomized controlled trial (RCT) of ovarian cancer screening (OCS) which is ongoing in the United Kingdom, involving 200,000 postmenopausal women (UK Collaborative Trial of Ovarian Cancer Screening, UKCTOCS, www.ukctocs.org) in a multimodality approach, estimates a woman's risk of ovarian cancer (ROC) [6]. Here, an estimate on the basis of age and modified by the relative fit of the serial CA125 profile to the change-point model estimated from known cases is compared with the flat profile model estimated from known controls. Women who are found to have a high ROC then undergo screening by transvaginal ultrasonography. Preliminary results are promising but it remains unknown if this multimodality screening with CA125 will improve disease specific mortality.

In order to improve detection, considerable research has been directed at furthering the understanding of the functional and biochemical nature of CA125. Three decades of research unveiled the partial nucleotide and amino acid sequence, cellular localization and secretion, oligosaccharide structures and possible biological roles of CA125 during cancer and normal conditions. Unfortunately, the presented data is incomplete and contradictory. In many cases it is difficult to draw conclusions from the reports.

This article summarizes the currently available published data and conflicting views on the molecular nature of CA125. As this ambiguity may arise from antibody cross-reactivity, an orthogonal method of protein identification is required. Several studies employ mass spectrometry (MS) to independently verify antibody based CA125 detection. However, currently published mass spectrometric data does not satisfy established protocols and could therefore be unreliable. This highlights the need for high quality mass spectrometric data to enable reliable CA125 identification. Ultimately, the revised knowledge about the nature of CA125 may lead to the development of quantification techniques that will increase current sensitivities and specificities of ovarian cancer diagnosis, prognosis and progression.

2. Biological Function

The biological role for CA125 is still under investigation though several studies have demonstrated a relationship between CA125 and the immune system. In 2003 Kui Wong *et al.* proposed a role in immunity based upon the characterization of the *N*-oligosaccharide structures identified on CA125 [7]. This hypothesis was supported by the study of Patankar *et al.* where natural killer (NK) cell function was found to be inhibited in the presence of CA125 purified from OVCAR-3 cell culture [8]. Cell membrane-bound CA125 was subsequently shown to bind directly to NK cells derived from peripheral blood of ovarian cancer patients and pregnant women [9]. A potential link between the suppression of NK cells in fetomaternal tolerance and the immune evasion of ovarian cancers was therefore proposed [9]. Further investigation into the mechanism of interaction revealed that binding was established through the sialic acid-binding Ig-like lectin-9 (Siglec-9), an inhibitory receptor expressed on NK cells [10,11].

Furthermore, the C-terminus of CA125 has been shown to interact specifically with galectin-1 which, on its own, has been shown to be upregulated in various cancer-derived cell lines [12] (reviewed in [13]). Studies have also reported CA125 binding to mesothelin, a glycosylphosphatidylinositol-linked cell surface protein [14]. It is hypothesized that this interaction may play a role in detachment, attachment and localized invasion that is unique to ovarian cancer metastasis [15]. In addition, CA125 was described as a calcium-dependent protease, as the addition of EDTA to FPLC-purified CA125 prevented auto-proteolysis [16]. However, the presence of other proteins (e.g., proteases) in the high molecular weight CA125 FPLC fraction was not investigated. To date, knowledge about many functional aspects of CA125 remains incomplete. Further investigations are therefore required to elucidate the role(s) of this molecule in pathological and physiological conditions.

3. Nucleotide and Amino Acid Sequence

The currently available genomic information on CA125 is based on two autonomous studies conducted in 2001 [17,18]. By screening a cDNA library generated using OVCAR-3 mRNA, Yin *et al.* identified a clone with a 5797 base pair (bp) nucleotide sequence [17] (reviewed in [19,20]). Although this region translated to a partial CA125 protein, it confirmed and revealed the existence of some key features. Importantly, they established the presence of partially conserved tandem repeats and were the first to describe the homology to the “sea urchin sperm protein, enterokinase, agrin” (SEA) domain. Yin *et al.* also postulated the existence of a transmembrane domain based upon a cluster of hydrophobic residues and a phosphorylation site based on a tyrosine phosphorylation consensus motif [17]. Furthermore, *O*-glycosylation was inferred due to the serine, threonine and proline content as well as *N*-glycosylation due to the asparagine content of the protein. These findings were supported and extended on by O’Brien and co-workers [18,21]. Here, a large transcript encoding CA125 was sequenced based upon a cyanogen bromide cleavage fragment of CA125 and a translated expressed sequence tag [16]. The resultant protein was described in three parts; an amino terminal domain, a repeat domain and a carboxy terminal domain (Table 1, Figure 1). The carboxy terminal domain reported by O’Brien *et al.* was homologous to the findings of Yin *et al.* [17,18,20]. One feature

described only by O'Brien was the existence of a potential proteolytic cleavage site at amino acid residues 171–181 [18].

Table 1. Features described from the CA125 amino acid sequence based on the data of O'Brien [18,21], Yin [17], Maeda [22] and Berman [23]; Amino acid (aa) number (numbering derived from [18]) indicates the position of those residues as they appear within the respective domain. bp: base pairs.

Domain Features	Carboxy Domain	Repeat Domain	Amino Domain
Exons	9	5	9
Nucleotides (bp)	14,000	Single: 1900 Cumulative: Unknown	50,950
Amino Acids (aa)	284	Single: 156 Cumulative: >9360	12,068
Molecular Characteristics	Stop Codon Poly A Signal Potential Transmembrane Region (aa 230–252) Potential Cytoplasmic Region (aa 256–260) Potential Proteolytic Cleavage Sites (aa 171–181 [18] and aa 35/36 [22])	SEA domain Cysteine loop (aa 59,79) [23] Conserved Methionine (aa 24) Serine/Threonine/Proline Rich	Serine/Threonine Rich (Abundant <i>O</i> -glycosylation Potential, <i>N</i> -glycosylation Potential)

As described previously, the repeat domain consists of partially conserved tandem repeats that are 156 amino acids in length. The SEA domain was noted to constitute the first 131 amino acids of this region [18]. However, the amino acid sequence exhibiting a SEA domain is variable in length as the main feature of the consensus sequence is a “strand-strand-helix-strand-strand-helix” secondary structure motif and is unaffected by small sequence gaps between these units [24]. Therefore, as most CA125 tandem repeats are polymorphic, the SEA domain may not necessarily be 131 amino acids long. This is exemplified by the identification of a SEA domain consensus sequence with a length of 59 amino acids located in the non-tandem repeat region of the C-terminal domain [17]. Furthermore, the repeat domain contains a highly conserved methionine at position 24 as well as two highly conserved cysteine residues at position 59 and 79 [18]. These cysteine residues in the repeat domain were hypothesized by O'Brien to form a 19 amino acid loop to which the M11 antibody binds. Additionally, O'Brien demonstrated a loss of binding of the M11 antibody to a recombinant repeat domain when digested with Asp-N or Lys-C as this would have disrupted the cysteine loop [18].

In the study by O'Brien *et al.*, 45 unique and 15 redundant tandem repeat sequences were identified [18]. It was concluded that CA125 consists of at least 60 tandem repeats. However, due to difficulties associated with sequencing repeat regions some may be wrongly positioned or unknown, leaving the exact number of repeats undefined [18]. Furthermore, O'Brien states that the genome

database is deficient in repeat information as repetitive sequences were potentially deleted during the compilation process of the human genome [18]. This suggests that the currently available nucleotide sequence of CA125 may not be correct.

The amino terminus was first described as a 1637 amino acid domain coded by 5 exons [17,18]. This was later revised by O’Brien *et al.* where an additional 37,700 nucleotides were sequenced from chromosome 19 [21]. Acquisition of the complete amino domain was confirmed by the identification of a Kozak translational initiation sequence as well as supporting Northern blot and mRNA expression data [21]. The main feature of this domain is the abundance of serine and threonine residues, which represent many potential sites of *O*-glycosylation, and disperse asparagine residues as potential sites of *N*-glycosylation [21].

Figure 1. Schematic of CA125. (A) A diagrammatic overview of the molecular structure of CA125 as described by O’Brien *et al.* [18,21]; (B) A typical repeat unit of 156 amino acids and the possible features assigned to this sequence. At least 60 repeats have been identified, of which 45 are unique sequences, giving rise to a minimum domain length of 9800 aa [18]. Serine/Threonine rich region downstream of SEA domain was unspecified [18], *O*-glycosylation potential of a randomly chosen repeat unit (aa 14,255–14,410) was determined by submission to NetOGlyc 3.1 Server [25] (aa 131–156); (C) Carboxy terminal domain: aa 21,869–22,152 (length: 284 aa) [18] (aa numbering according to UniProt [26]).

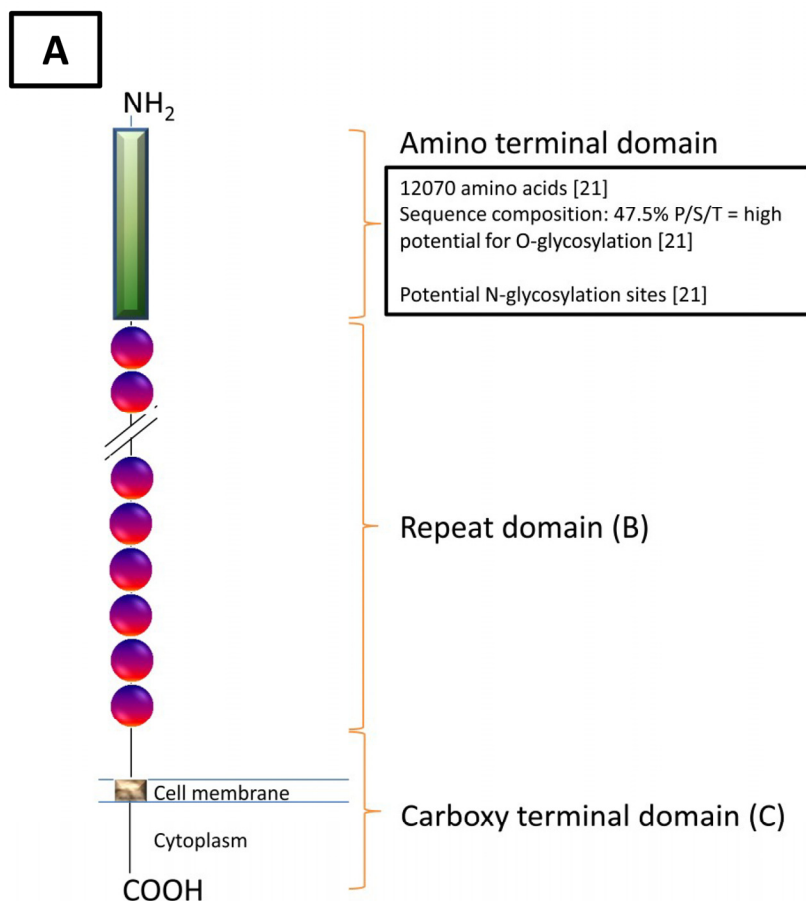
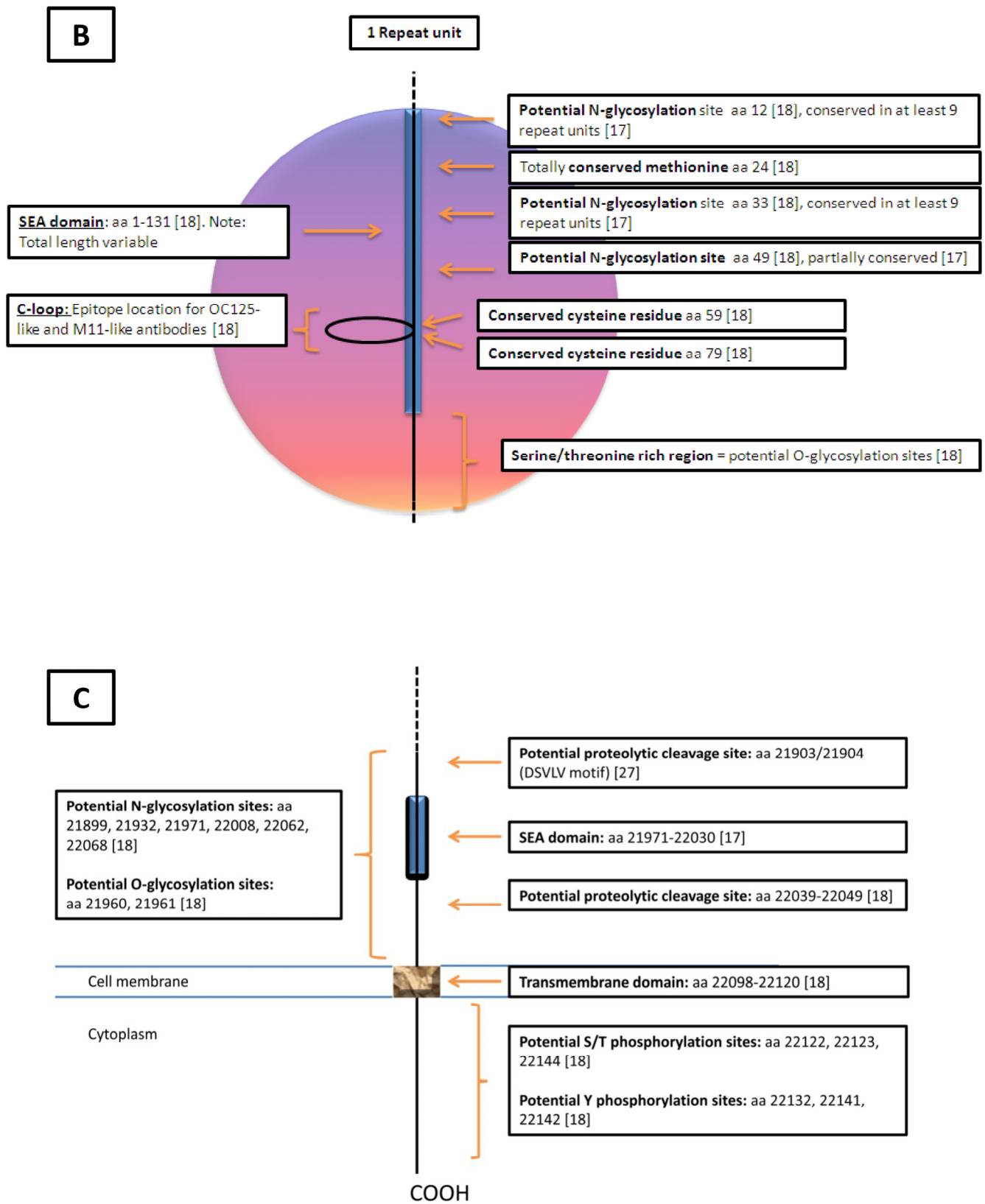


Figure 1. Cont.



The elucidation of the almost complete CA125 nucleotide sequence represented a significant leap in understanding. It enabled clarification of known and/or suspected features such as glycosylation (amino domain), multivalent antibody binding (repeat domain) as well as anchorage to and release

from the membrane (carboxy domain) [12,27]. The findings also raise questions regarding the generation of lower molecular weight isoforms. Only the existence of a CA125 precursor protein of approximately 400 kDa has been found to exist in the cytoplasm of OVCAR-3 cells prior to protein maturation [28]. CA125 variants are also thought to arise from proteolytic cleavage. Proteins such as MUC1 have a proteolytic cleavage motif, GSVVV, in their SEA domain [29,30], however the presence of a motif with a similar consensus sequence has only been confirmed in one SEA domain of CA125 thus far [22]. A cluster analysis of a human CA125 sequence, containing the first 12 tandem repeats and a murine homolog to the C-terminus of CA125, was performed by Maeda *et al.* [22]. Interestingly, a DSVLV sequence in the second SEA domain from the cytosolic side of CA125 was clustered with the GSVVV sequence. Therefore, it was concluded that proteolytic cleavage might also occur in this domain of CA125 [22]. However, the formation of further proteolytic fragments of CA125, by the mechanism of the DSVLV or GSVVV sequence, in additional SEA domains has not been reported. Furthermore, O'Brien *et al.* suggested that the formation of alternative splicing products of this protein seems unlikely, based on the inherent variation of the five exons encoding for each repeat sequence [18]. Therefore, a discrepancy exists between antibody findings of low molecular weight isoforms and the protein sequence data (discussed below).

4. Antigenic Determinants

Numerous antibodies have been generated against CA125 since the initial discovery in 1981. These antibodies are classified as OC125-like (group A), M11-like (group B) or OV197 (group C) depending on the antigenic determinant recognized [31]. These antibodies are further subdivided into groups A1–A4 (OC125-like), B1/B2 (M11-like) and C1/C2 (OV197-like) [32,33]. It is important to note that although there are two antibodies classified as OV197-like, the antigenic recognition behavior of OV197 is different to that of 7C12 [32]. Extensive investigation into the epitope recognition of the different class antibodies to CA125 has been performed by the International Society of Oncology and Biomarkers (ISOBM) TD-1 workshop [31]. Cross-inhibition and immunometric assays have been used to examine the various epitopes on CA125 isolated from normal abdominal fluids, cervical mucus, cell culture and ascites [32]. From this the importance of antibody class combinations for accurate detection of CA125 was revealed. Furthermore, differential antibody interaction with CA125 isolated in low-molecular-weight fractions was also highlighted. Here, different immunometric assay combinations resulted in different CA125 activity levels in those fractions [32]. It was noted that the difference in assay results could be due to the different behavior of the antibody combinations towards the low-molecular-weight fractions of CA125 preparations [32]. Conversely, fractions containing high-molecular-weight CA125 yielded similar CA125 activity levels regardless of the antibody pair used in the immune-assay. As high-molecular-weight forms of CA125 are the major component in most samples this phenomenon should, therefore, not interfere with the immunometric assays even if the standards used do not have the same composition as the sample [32].

OC125-like and M11-like antibodies are proposed to detect a cysteine enclosed loop (C-loop) structure in the repeat domain [18,31,32] with a paucity of glyco-structures in the immediate proximity [18]. Extrapolated from the C-loop of a murine derived CA125 repeat domain homolog, Maeda *et al.* suggested that this antigenic region forms a β -sheet structure [22]. This was later

confirmed with a synthesized 21 amino acid long human CA125 antigenic domain [23]. Here, a serine in position 8 of the loop was found to be essential for structural formation [23]. However, it is important to note that this serine at position 8 is conserved in the repeat sequences, throughout human CA125, in only around 25% of the cases. In approximately 60% of repeat sequences a proline is found at position 8, which results in random coil formation [23].

The fourth report from ISOBM TD-1 workshop described the use of a recombinant sequence of the 11th repeat from the *N*-terminus (R11) to test the binding capacity of various OC125-like, M11-like and OV197-like antibodies [33]. It was shown that the binding of the various monoclonal antibodies differed between the groups and subgroups. Within group A, subgroups A3 and A4 showed high binding to R11. Group B also displayed high reactivity with the exception of four investigated M11-like antibodies that showed low to no binding [33]. Additionally, group C antibodies had shown only little or no reactivity [33]. From those results it was concluded that the poor or absent binding of some antibodies towards R11 is unlikely to be due to a disparate sequence of this particular repeat to the other +60 repeat sequences, as there are only three main epitope families identified to date [33]. This observation was thought to be, more likely, a result of the single R11 repeat not exhibiting an optimal conformation as it is no longer in its native context [33]. This was exemplified by the binding behavior of OC125, which binds poorly to R11 [33] yet shows strong binding to a recombinant 3 tandem repeat form of CA125 [34].

Furthermore, glycosylation of CA125 has also been shown to be important for high affinity antibody binding. Deglycosylation with PNGase F led to lower affinity binding of the OC125 antibody to CA125 [7]. However, the exact influence glycosylation has on the binding affinity of the different antibody classes to CA125 is unknown.

5. Forms and Variants

Several different masses corresponding to CA125 have been reported in the literature. A protein with a mass over 1000 kDa as well as proteins with a lower molecular mass ranging from 200–400 kDa, 130 kDa, 205 kDa and 55 kDa have been observed depending on the source material analyzed [35–37] (reviewed in [38]). A number of these studies have also reported a modular structure of CA125, which can be broken down into 200 kDa, 70 kDa and 50 kDa species [35,37]. These forms of CA125 have been referred to as splice variants, which give rise to amino acid sequences with a molecular mass between 49 and 1500 kDa [22,39].

However, studies by Lloyd and Yin found no evidence for the low molecular weight products reported by other groups using an OVCAR-3 cell line [28,40]. Later they reported a precursor form of CA125 of approximately 400 kDa in an OVCAR-3 cell line, from which the mature form of the protein becomes assembled [28]. It is noted that only the mature protein form is released into the surrounding medium [28]. Additionally, the low molecular weight forms reported by other groups were assessed as possible degradation products by using trypsin [28]. This resulted in no defined, prominent molecular weight bands on a SDS-PAGE gel [28]. Furthermore, our research has not found evidence of low molecular weight forms of CA125 in human ascites [41]. Here, cross-reactivity of the employed M11-like and OC125-like antibodies was shown by two-dimensional western-blotting, mass

spectrometric identification and spiking experiments. Therefore, the existence of lower molecular weight forms of CA125 remains uncertain.

6. Oligosaccharide Moieties

Mucins such as CA125 are large proteins associated with extensive glycosylation that is important for structure and function (reviewed in [42]). In addition to the contradicting views regarding the molecular weight forms of CA125, the types of glycosylation reported are also controversial. Davis *et al.* identified CA125 as protein with, in comparison to other mucins, low carbohydrate content of 24%, constituting mostly mannose and N-linked glycans [35], while Hanisch *et al.* described CA125 as consisting of 68% carbohydrates, essentially mannose free and O-linked glycans [43].

A third view was presented by Nagata *et al.* in 1991 [44]. They found CA125 to be a glycosylphosphatidylinositol (GPI) anchored molecule with N- and O-linked glycan structures [44]. Independently, a study by Kui Wong *et al.* also found N- and O-linked glycans attached to CA125 [7]. Interestingly, one identified structure was identical to an unusual glycostructure expressed by uromodulin [7,45]. In light of our recent results these contradicting reports may arise from cross-reactivity of the M11-like and OC125-like antibodies with other proteins [41]. This may have lead to the characterization of proteins that were wrongly assigned as CA125.

Additionally, aberrant glycosylation of CA125 has been reported for various diseases. The study of Pastan *et al.* demonstrated increased presentation of Lewis y blood group tetrasaccharide (Le^y) on ovarian cancer tissue compared with normal ovarian tissue [46]. This Le^y modification, expressed in 75% of ovarian cancer tissue specimens, was identified as a part of the CA125 glycan structure [47]. Microheterogeneities in CA125 glycosylation may represent a possible diagnostic indicator, according to Jankovic *et al.*, as differential glycosylation of CA125 derived from amniotic fluid and OVCAR-3 cells was demonstrated [48]. Furthermore, Mitic *et al.* reported differential binding of CA125 to a set of sialic acid-binding Ig-like lectins (Siglecs) depending on the source of the protein. A preference for Siglec-9 binding was evident for CA125 derived from an OVCAR-3 cell line [49]. Interestingly, this interaction has been shown to be necessary for the binding to and subsequent suppression of NK-cell function [8–11].

7. Mass Spectrometric Identification of CA125

Isolation of CA125 for the analysis of its molecular features represents a challenge for researchers, as routine high resolution methods such as SDS-PAGE are not ideal for very high molecular weight analytes. Typically, size exclusion chromatography (SEC) is used for CA125 purification; however the separation efficiency is limited and results in cross-contaminations [16]. Additional methods used for the prefractionation of CA125 include immuno-precipitation, affinity chromatography or a combination of these methods that are associated with similar disadvantages [12,16,28,48,50]. Subsequent identification of CA125 in the derived fractions is therefore necessary for reliable downstream analysis. However, CA125 identification solely based on antibody probing alone can lead to false-positive protein assignment as a result of antibody cross-reactivity [41]. Thus, characterization of these secondary proteins may lead to erroneous assignment of molecular attributes to CA125. Several studies show that mass spectrometry (MS) offers a suitable solution for independent CA125

identification (discussed below). It is, however, notable that the published protein sequence for CA125 is incomplete [17,18,21]. Although this may result in some peptide identifications being missed by search algorithms, mass spectrometric identification of CA125 is a viable approach that can complement antibody based identification methods.

The first peptide mass fingerprinting data on CA125 was described in 2003 where 16 masses, corresponding to peptides of a 1148 amino acid form of CA125 from HeLa cells, were detected [12]. This enabled the identification of CA125 as a binding protein of galectin-1. In a later study, Jankovic *et al.* identified an approximately 200 kDa form of CA125 in first trimester human placental extract based on 36 experimentally detected masses [51]. Similarly, Milutinovic *et al.* reported identification of 46 masses that matched to CA125 in human amniotic fluid [52].

Tandem mass spectrometric data on CA125 was first reported by Kui Wong *et al.*, confirming the identity of CA125 in a purified sample prior to oligosaccharide characterization [7]. Although only a single peptide was reported, this identification was supported by peptide mass fingerprinting where 11 experimentally derived masses were matched to theoretical peptides. Analyses of the tear proteome by tandem MS lead to the identification of 491 proteins including CA125 [53]. This was the first mass spectrometric data of CA125 that included statistical analysis and associated metadata for eight peptides of this protein. However, the molecular weight range of the fraction in which CA125 was identified was not stated. Using Fourier-transform ion cyclotron resonance (FT-ICR) tandem mass spectrometry Andersch-Bjorkman *et al.* detected 31 peptides from a 2–3 MDa form of CA125 derived from mucus of the human cervical opening [54]. Davies *et al.* also reported the identification of a high molecular weight form of CA125 in human tracheal secretions [55]. Recently, we have also presented mass spectrometric data including statistical analysis and metadata of 21 peptides of a high molecular weight species of CA125 [41]. These data were used to confirm the identity of CA125 in the ascites of ovarian cancer patients as a high molecular weight protein.

To date, mass spectrometric data of CA125 is still quite rare and seldom reported in a way that satisfies the “Minimum Information about a Proteomics Experiment” (MIAPE) standards [56–58], even after 2007. MIAPE standards have been implemented to avoid confusion about mass spectrometric experiments as well as enhance reproducibility and reliability of MS data. It is evident that the quality of mass spectrometric data regarding CA125 varies greatly and the majority of studies do not allow for independent assessment. Furthermore, statistical analysis of the MS-generated data is seldom presented. This is a major issue, especially for CA125, as the potential for incorrect peptide matches rises with increasing length of protein sequence [59].

Additionally, matching experimentally found masses to theoretical masses without the usage of statistical tools is prone to produce false-positive identifications. Schober *et al.* demonstrated that when a mass accuracy window greater than ± 10 ppm is used, numerous peptides from different proteins with similar masses can be assigned to the incorrect protein [60]. This means that CA125 identification without supporting statistical analysis is unreliable and potentially false-positive.

8. Conclusion

CA125 is heavily relied upon for diagnosis and prognosis of epithelial ovarian cancer; however the molecular characteristics are poorly understood and are the subject of ongoing investigations. This

review summarizes and evaluates the available data regarding the molecular features of CA125. While the nucleotide sequence of this protein was independently cloned by two groups and showed an almost complete homology between the overlapping parts, other molecular features like predominant type of glycosylation and isoforms are not only disputed but mutually exclusive. Although mass spectrometry has already been shown to be a valid option to clarify these ambiguous findings, most data generated by this method, and published so far, does not satisfy established reporting standards. This review highlights the importance of high quality mass spectrometric data to independently verify the identity of the investigated protein. This would ensure that CA125 identified by antibody-dependent methods are not cross-reactive proteins. Therefore, the ambiguous views on the glycan structure as well as the controversial low molecular weight forms can be clarified. This greater understanding may help improve established measurement techniques for assaying CA125 serum levels. Alternatively, novel detection systems that are independent of antibodies, such as selective reaction monitoring mass spectrometry (SRM), could also be applied for CA125 quantitation in patient samples. SRM has already been shown to be a feasible approach to quantify proteins in plasma [61]. Furthermore, detailed structural information on CA125 will enable the development of new and more refined molecularly targeted antibody therapies which already show great clinical potential and warrant further research [62–65].

Acknowledgments

This study was supported by the Ovarian Cancer Research Foundation.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Bast, R.C., Jr.; Feeney, M.; Lazarus, H.; Nadler, L.M.; Colvin, R.B.; Knapp, R.C. Reactivity of a monoclonal antibody with human ovarian carcinoma. *J. Clin. Invest.* **1981**, *68*, 1331–1317.
2. Bast, R.C., Jr.; Klug, T.L.; St John, E.; Jenison, E.; Niloff, J.M.; Lazarus, H.; Berkowitz, R.S.; Leavitt, T.; Griffiths, C.T.; Parker, L.; *et al.* A radioimmunoassay using a monoclonal antibody to monitor the course of epithelial ovarian cancer. *N. Engl. J. Med.* **1983**, *309*, 883–887.
3. Muyldermans, M.; Cornillie, F.J.; Koninckx, P.R. CA125 and endometriosis. *Hum. Reprod. Update* **1995**, *1*, 173–187.
4. Jacobs, I.; Oram, D.; Fairbanks, J.; Turner, J.; Frost, C.; Grudzinskas, J.G. A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. *Br. J. Obstet. Gynaecol.* **1990**, *97*, 922–929.
5. Bast, R.C., Jr.; Badgwell, D.; Lu, Z.; Marquez, R.; Rosen, D.; Liu, J.; Baggerly, K.A.; Atkinson, E.N.; Skates, S.; Zhang, Z.; *et al.* New tumor markers: CA125 and beyond. *Int. J. Gynecol. Cancer* **2005**, *15 Suppl 3*, S274–S281.

6. Menon, U.; Gentry-Maharaj, A.; Hallett, R.; Ryan, A.; Burnell, M.; Sharma, A.; Lewis, S.; Davies, S.; Philpott, S.; Lopes, A.; *et al.* Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: Results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *Lancet Oncol.* **2009**, *10*, 327–340.
7. Kui Wong, N.; Easton, R.L.; Panico, M.; Sutton-Smith, M.; Morrison, J.C.; Lattanzio, F.A.; Morris, H.R.; Clark, G.F.; Dell, A.; Patankar, M.S. Characterization of the oligosaccharides associated with the human ovarian tumor marker CA125. *J. Biol. Chem.* **2003**, *278*, 28619–28634.
8. Patankar, M.S.; Jing, Y.; Morrison, J.C.; Belisle, J.A.; Lattanzio, F.A.; Deng, Y.; Wong, N.K.; Morris, H.R.; Dell, A.; Clark, G.F. Potent suppression of natural killer cell response mediated by the ovarian tumor marker CA125. *Gynecol. Oncol.* **2005**, *99*, 704–713.
9. Belisle, J.A.; Gubbels, J.A.; Raphael, C.A.; Migneault, M.; Rancourt, C.; Connor, J.P.; Patankar, M.S. Peritoneal natural killer cells from epithelial ovarian cancer patients show an altered phenotype and bind to the tumour marker MUC16 (CA125). *Immunology* **2007**, *122*, 418–429.
10. Zhang, J.Q.; Nicoll, G.; Jones, C.; Crocker, P.R. Siglec-9, a novel sialic acid binding member of the immunoglobulin superfamily expressed broadly on human blood leukocytes. *J. Biol. Chem.* **2000**, *275*, 22121–22126.
11. Belisle, J.A.; Horibata, S.; Jennifer, G.A.; Petrie, S.; Kapur, A.; Andre, S.; Gabius, H.J.; Rancourt, C.; Connor, J.; Paulson, J.C.; *et al.* Identification of Siglec-9 as the receptor for MUC16 on human NK cells, B cells, and monocytes. *Mol. Cancer* **2010**, *9*, 118.
12. Seelenmeyer, C.; Wegehngel, S.; Lechner, J.; Nickel, W. The cancer antigen CA125 represents a novel counter receptor for galectin-1. *J. Cell Sci.* **2003**, *116* (Pt 7), 1305–1318.
13. Demydenko, D.; Berest, I. Expression of galectin-1 in malignant tumors. *Exp. Oncol.* **2009**, *31*, 74–79.
14. Rump, A.; Morikawa, Y.; Tanaka, M.; Minami, S.; Umesaki, N.; Takeuchi, M.; Miyajima, A. Binding of ovarian cancer antigen CA125/MUC16 to mesothelin mediates cell adhesion. *J. Biol. Chem.* **2004**, *279*, 9190–9198.
15. Rodriguez, G.C.; Haisley, C.; Hurteau, J.; Moser, T.L.; Whitaker, R.; Bast, R.C., Jr.; Stack, M.S. Regulation of invasion of epithelial ovarian cancer by transforming growth factor-beta. *Gynecol. Oncol.* **2001**, *80*, 245–253.
16. O'Brien, T.J.; Tanimoto, H.; Konishi, I.; Gee, M. More than 15 years of CA 125: What is known about the antigen, its structure and its function. *Int. J. Biol. Markers* **1998**, *13*, 188–195.
17. Yin, B.W.; Lloyd, K.O. Molecular cloning of the CA125 ovarian cancer antigen: Identification as a new mucin, MUC16. *J. Biol. Chem.* **2001**, *276*, 27371–27375.
18. O'Brien, T.J.; Beard, J.B.; Underwood, L.J.; Dennis, R.A.; Santin, A.D.; York, L. The CA 125 gene: An extracellular superstructure dominated by repeat sequences. *Tumour Biol.* **2001**, *22*, 348–366.
19. McLemore, M.R.; Aouizerat, B. Introducing the MUC16 gene: Implications for prevention and early detection in epithelial ovarian cancer. *Biol. Res. Nurs.* **2005**, *6*, 262–267.
20. Hovig, E.; Rye, P.D.; Warren, D.J.; Nustad, K. CA 125: The end of the beginning. *Tumour Biol.* **2001**, *22*, 345–347.

21. O'Brien, T.J.; Beard, J.B.; Underwood, L.J.; Shigemasa, K. The CA 125 gene: A newly discovered extension of the glycosylated *N*-terminal domain doubles the size of this extracellular superstructure. *Tumour Biol.* **2002**, *23*, 154–169.
22. Maeda, T.; Inoue, M.; Koshiha, S.; Yabuki, T.; Aoki, M.; Nunokawa, E.; Seki, E.; Matsuda, T.; Motoda, Y.; Kobayashi, A.; *et al.* Solution structure of the SEA domain from the murine homologue of ovarian cancer antigen CA125 (MUC16). *J. Biol. Chem.* **2004**, *279*, 13174–13182.
23. Berman, Z.T.; Moore, L.J.; Knudson, K.E.; Whelan, R.J. Synthesis and structural characterization of the peptide epitope of the ovarian cancer biomarker CA125 (MUC16). *Tumour Biol.* **2010**, *31*, 495–502.
24. Bork, P.; Patthy, L. The SEA module: A new extracellular domain associated with *O*-glycosylation. *Protein Sci.* **1995**, *4*, 1421–1425.
25. Julenius, K.; Molgaard, A.; Gupta, R.; Brunak, S. Prediction, conservation analysis, and structural characterization of mammalian mucin-type *O*-glycosylation sites. *Glycobiology* **2005**, *15*, 153–164.
26. Apweiler, R.; Bairoch, A.; Wu, C.H.; Barker, W.C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; Magrane, M.; *et al.* UniProt: The Universal Protein knowledgebase. *Nucleic Acids Res.* **2004**, *32* (Database issue), D115–D119.
27. Konishi, I.; Fendrick, J.L.; Parmley, T.H.; Quirk, J.G., Jr.; O'Brien, T.J. Epidermal growth factor enhances secretion of the ovarian tumor-associated cancer antigen CA125 from the human amnion WISH cell line. *J. Soc. Gynecol. Investig.* **1994**, *1*, 89–96.
28. Lloyd, K.O.; Yin, B.W. Synthesis and secretion of the ovarian cancer antigen CA 125 by the human cancer cell line NIH:OVCAR-3. *Tumour Biol.* **2001**, *22*, 77–82.
29. Parry, S.; Silverman, H.S.; McDermott, K.; Willis, A.; Hollingsworth, M.A.; Harris, A. Identification of MUC1 proteolytic cleavage sites *in vivo*. *Biochem. Biophys. Res. Commun.* **2001**, *283*, 715–720.
30. Palmai-Pallag, T.; Khodabukus, N.; Kinarsky, L.; Leir, S.H.; Sherman, S.; Hollingsworth, M.A.; Harris, A. The role of the SEA (sea urchin sperm protein, enterokinase and agrin) module in cleavage of membrane-tethered mucins. *FEBS J.* **2005**, *272*, 2901–2911.
31. Nustad, K.; Bast, R.C., Jr.; Brien, T.J.; Nilsson, O.; Seguin, P.; Suresh, M.R.; Saga, T.; Nozawa, S.; Borner, O.P.; de Bruijn, H.W.; *et al.* Specificity and affinity of 26 monoclonal antibodies against the CA 125 antigen: First report from the ISOBM TD-1 workshop. International Society for Oncodevelopmental Biology and Medicine. *Tumour Biol.* **1996**, *17*, 196–219.
32. Nustad, K.; Lebedin, Y.; Lloyd, K.O.; Shigemasa, K.; de Bruijn, H.W.; Jansson, B.; Nilsson, O.; Olsen, K.H.; O'Brien, T.J. Epitopes on CA 125 from cervical mucus and ascites fluid and characterization of six new antibodies. Third report from the ISOBM TD-1 workshop. *Tumour Biol.* **2002**, *23*, 303–314.
33. Warren, D.J.; Nustad, K.; Beard, J.B.; O'Brien, T.J. Expression and epitope characterization of a recombinant CA 125 repeat: Fourth report from the ISOBM TD-1 workshop. *Tumour Biol.* **2009**, *30*, 51–60.
34. Yin, B.W.; Dnistrian, A.; Lloyd, K.O. Ovarian cancer antigen CA125 is encoded by the MUC16 mucin gene. *Int. J. Cancer* **2002**, *98*, 737–740.
35. Davis, H.M.; Zurawski, V.R., Jr.; Bast, R.C., Jr.; Klug, T.L. Characterization of the CA 125 antigen associated with human epithelial ovarian carcinomas. *Cancer Res.* **1986**, *46* (12 Pt 1), 6143–6148.

36. Halila, H. Detection of ovarian cancer marker CA 125 in human seminal plasma. *Tumour Biol.* **1985**, *6*, 207–212.
37. De los Frailes, M.T.; Stark, S.; Jaeger, W.; Hoerauf, A.; Wildt, L. Purification and characterization of the CA 125 tumor-associated antigen from human ascites. *Tumour Biol.* **1993**, *14*, 18–29.
38. Bouanene, H.; Miled, A. Conflicting views on the molecular structure of the cancer antigen CA125/MUC16. *Dis. Markers* **2010**, *28*, 385–394.
39. Grimwood, J.; Gordon, L.A.; Olsen, A.; Terry, A.; Schmutz, J.; Lamerdin, J.; Hellsten, U.; Goodstein, D.; Couronne, O.; Tran-Gyamfi, M.; *et al.* The DNA sequence and biology of human chromosome 19. *Nature* **2004**, *428*, 529–535.
40. Lloyd, K.O.; Yin, B.W.; Kudryashov, V. 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* **1997**, *71*, 842–850.
41. Weiland, F.; Fritz, K.; Oehler, M.K.; Hoffmann, P. Methods for identification of CA125 from ovarian cancer ascites by high resolution mass spectrometry. *Int. J. Mol. Sci.* **2012**, *13*, 9942–9958.
42. Tian, E.; Ten Hagen, K.G. Recent insights into the biological roles of mucin-type O-glycosylation. *Glycoconj. J.* **2009**, *26*, 325–334.
43. Hanisch, F.G.; Uhlenbruck, G.; Dienst, C.; Stottrop, M.; Hippauf, E. Ca 125 and Ca 19-9: Two cancer-associated sialylsaccharide antigens on a mucus glycoprotein from human milk. *Eur. J. Biochem.* **1985**, *149*, 323–330.
44. Nagata, A.; Hirota, N.; Sakai, T.; Fujimoto, M.; Komoda, T. Molecular nature and possible presence of a membranous glycan-phosphatidylinositol anchor of CA125 antigen. *Tumour Biol.* **1991**, *12*, 279–286.
45. Easton, R.L.; Patankar, M.S.; Clark, G.F.; Morris, H.R.; Dell, A. Pregnancy-associated changes in the glycosylation of tamm-horsfall glycoprotein. Expression of sialyl Lewis(x) sequences on core 2 type O-glycans derived from uromodulin. *J. Biol. Chem.* **2000**, *275*, 21928–21938.
46. Pastan, I.; Lovelace, E.T.; Gallo, M.G.; Rutherford, A.V.; Magnani, J.L.; Willingham, M.C. Characterization of monoclonal antibodies B1 and B3 that react with mucinous adenocarcinomas. *Cancer Res.* **1991**, *51*, 3781–3787.
47. Yin, B.W.; Finstad, C.L.; Kitamura, K.; Federici, M.G.; Welshinger, M.; Kudryashov, V.; Hoskins, W.J.; Welt, S.; Lloyd, K.O. Serological and immunochemical analysis of Lewis y (Ley) blood group antigen expression in epithelial ovarian cancer. *Int. J. Cancer* **1996**, *65*, 406–412.
48. Jankovic, M.M.; Milutinovic, B.S. Glycoforms of CA125 antigen as a possible cancer marker. *Cancer Biomark.* **2008**, *4*, 35–42.
49. Mitic, N.; Milutinovic, B.; Jankovic, M. Assessment of sialic acid diversity in cancer- and non-cancer related CA125 antigen using sialic acid-binding Ig-like lectins (Siglecs). *Dis. Markers* **2012**, *32*, 187–194.
50. Bouanene, H.; Saibi, W.; Mokni, M.; Sriha, B.; Ben Fatma, L.; Ben Limem, H.; Ben Ahmed, S.; Gargouri, A.; Miled, A. Biochemical and morphological differences between CA125 isolated from healthy women and patients with epithelial ovarian cancer from Tunisian population. *Pathol. Oncol. Res.* **2012**, *18*, 325–330.
51. Jankovic, M.M.; Tapuskovic, B.S. Molecular forms and microheterogeneity of the oligosaccharide chains of pregnancy-associated CA125 antigen. *Hum. Reprod.* **2005**, *20*, 2632–2638.

52. Milutinovic, B.; Jankovic, M. Analysis of the protein and glycan parts of CA125 Antigen from human amniotic fluid. *Arch. Biol. Sci. Belgrade* **2007**, *59*, 97–103.
53. De Souza, G.A.; Godoy, L.M.; Mann, M. Identification of 491 proteins in the tear fluid proteome reveals a large number of proteases and protease inhibitors. *Genome Biol.* **2006**, *7*, R72.
54. Andersch-Bjorkman, Y.; Thomsson, K.A.; Holmen Larsson, J.M.; Ekerhovd, E.; Hansson, G.C. Large scale identification of proteins, mucins, and their O-glycosylation in the endocervical mucus during the menstrual cycle. *Mol. Cell. Proteomics* **2007**, *6*, 708–716.
55. Davies, J.R.; Kirkham, S.; Svitacheva, N.; Thornton, D.J.; Carlstedt, I. 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 Biol.* **2007**, *39*, 1943–1954.
56. Taylor, C.F.; Paton, N.W.; Lilley, K.S.; Binz, P.A.; Julian, R.K., Jr.; Jones, A.R.; Zhu, W.; Apweiler, R.; Aebersold, R.; Deutsch, E.W.; *et al.* The minimum information about a proteomics experiment (MIAPE). *Nat. Biotechnol.* **2007**, *25*, 887–893.
57. Taylor, C.F.; Binz, P.A.; Aebersold, R.; Affolter, M.; Barkovich, R.; Deutsch, E.W.; Horn, D.M.; Huhmer, A.; Kussmann, M.; Lilley, K.; *et al.* Guidelines for reporting the use of mass spectrometry in proteomics. *Nat. Biotechnol.* **2008**, *26*, 860–861.
58. Binz, P.A.; Barkovich, R.; Beavis, R.C.; Creasy, D.; Horn, D.M.; Julian, R.K., Jr.; Seymour, S.L.; Taylor, C.F.; Vandenbrouck, Y. Guidelines for reporting the use of mass spectrometry informatics in proteomics. *Nat. Biotechnol.* **2008**, *26*, 862.
59. States, D.J.; Omenn, G.S.; Blackwell, T.W.; Fermin, D.; Eng, J.; Speicher, D.W.; Hanash, S.M. Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. *Nat. Biotechnol.* **2006**, *24*, 333–338.
60. Schober, Y.; Schramm, T.; Spengler, B.; Rompp, A. Protein identification by accurate mass matrix-assisted laser desorption/ionization imaging of tryptic peptides. *Rapid. Commun. Mass Spectrom.* **2011**, *25*, 2475–2483.
61. Ji, Q.C.; Rodila, R.; Gage, E.M.; El-Shourbagy, T.A. A strategy of plasma protein quantitation by selective reaction monitoring of an intact protein. *Anal. Chem.* **2003**, *75*, 7008–7014.
62. Chen, Y.; Clark, S.; Wong, T.; Dennis, M.S.; Luis, E.; Zhong, F.; Bheddah, S.; Koeppen, H.; Gogineni, A.; Ross, S.; *et al.* Armed antibodies targeting the mucin repeats of the ovarian cancer antigen, MUC16, are highly efficacious in animal tumor models. *Cancer Res.* **2007**, *67*, 4924–4932.
63. Reinartz, S.; Kohler, S.; Schlebusch, H.; Krista, K.; Giffels, P.; Renke, K.; Huober, J.; Mobus, V.; Kreienberg, R.; DuBois, A.; *et al.* Vaccination of patients with advanced ovarian carcinoma with the anti-idiotypic ACA125: Immunological response and survival (phase Ib/II). *Clin. Cancer Res.* **2004**, *10*, 1580–1587.
64. Leffers, N.; Daemen, T.; Helfrich, W.; Boezen, H.M.; Cohlen, B.J.; Melief, K.; Nijman, H.W. Antigen-specific active immunotherapy for ovarian cancer. *Cochrane Database Syst. Rev.* **2010**, *1*, CD007287.
65. Oei, A.L.; Sweep, F.C.; Thomas, C.M.; Boerman, O.C.; Massuger, L.F. The use of monoclonal antibodies for the treatment of epithelial ovarian cancer (review). *Int. J. Oncol.* **2008**, *32*, 1145–1157.