ScFv multimers of the anti-neuraminidase antibody NC10: shortening of the linker in single-chain Fv fragment assembled in \( \text{V}_L \) to \( \text{V}_H \) orientation drives the formation of dimers, trimers, tetrarsers and higher molecular mass multimers

Olan Dolezal\(^1\), Lesley A. Pearce, Lynne J. Lawrence\(^2\), Airlie J. McCoy\(^3\), Peter J. Hudson and Alexander A. Kortt

CSIRO Health Sciences and Nutrition and CRC for Diagnostic Technologies, 343 Royal Parade, Parkville, Victoria, Australia 3052 and
\(^1\)Department of Haematology, Cambridge University, Hills Road, Cambridge CB2 2XY, UK

To whom correspondence should be addressed.
E-mail: olan.dolezal@hsn.csiro.au

Synthetic genes encoding single-chain variable fragments (scFvs) of NC10 anti-neuraminidase antibody were constructed by joining the \( \text{V}_L \) and \( \text{V}_H \) domains with linkers of fifteen, five, four, three, two, one and zero residues. These \( \text{V}_L–\text{V}_H \) constructs were expressed in \textit{Escherichia coli} and the resulting proteins were characterized and compared with the previously characterized NC10 scFv proteins assembled in \( \text{V}_H–\text{V}_L \) orientation. Size-exclusion chromatography and electron microscope images of complexes formed between various NC10 scFvs and anti-idiotypic Fab were used to analyse the oligomeric status of these scFvs. The result showed that as the linker length between \( \text{V}_L \) and \( \text{V}_H \) was reduced, different patterns of oligomerization were observed compared with those with \( \text{V}_H–\text{V}_L \) isomers. As was the case for \( \text{V}_H–\text{V}_L \) orientation, the scFv-15 \( \text{V}_L–\text{V}_H \) protein existed mainly as a monomer whereas dimer (diabody) was a predominant conformation for the scFv-5, scFv-4 and scFv-3 \( \text{V}_L–\text{V}_H \) proteins. In contrast to the \( \text{V}_H–\text{V}_L \) isomer, direct ligation of \( \text{V}_L \) to \( \text{V}_H \) led to the formation of predominantly a tetramer (tetrabody) rather than to an expected trimer (triabody). Furthermore, the transition between dimers and higher order oligomers was not as distinct as for \( \text{V}_H–\text{V}_L \). Thus reducing the linker length in \( \text{V}_L–\text{V}_H \) from three to two residues did not precisely dictate a transition between dimers and trimers. Instead, two-residue as well as one-residue linked scFvs formed a mixture of dimers, trimers and tetrayers.

Keywords: antibody/diabody/single-chain Fvs/tetrabody/triabody

Introduction

The antigen-binding portion of an antibody molecule (Fv fragment) is formed by the association of the heavy chain variable region (\( \text{V}_H \)) and the light chain variable region (\( \text{V}_L \)). Fv fragments are the smallest entities that consistently maintain the binding specificity of the whole antibody. Although recombinant DNA techniques have facilitated individual Fv domain production, the non-covalently associated \( \text{V}_L \) and \( \text{V}_H \) domains in an individual Fv fragment tend to dissociate from one another (Glockshuber et al., 1990). To improve stability, recombinant single chain Fv fragments (scFvs) have been engineered with two variable domains covalently joined via a flexible peptide linker (Bird et al., 1988; Huston et al., 1988). In general, scFvs with linkers >12 amino acid residues in length form stable monomers, which usually exhibit similar antigen binding affinity to the parent antibody (Skerra et al., 1991; Kortt et al., 1994). However, scFv monomers have also been observed to form active dimers and higher molecular mass multimers upon freezing at high protein concentrations (Kortt et al., 1994).

To provide increased avidity to target antigens, recent attention has focused on the design of linkers which generate multivalent scFv molecules (reviewed by Hudson, 1999). Multivalent scFv molecules are sufficiently large to avoid the fast clearance through the kidneys observed for scFv monomers and thereby have potential application for tumour imaging and radiotherapy (Adams et al., 1998; Colcher et al., 1998; Wu et al., 1999). Construction of stable multimeric scFv molecules can be achieved by the shortening of the linker length to <12 residues such that the \( \text{V}_L \) domain is unable to associate with its attached \( \text{V}_H \) domain and thus generate a monomeric Fv fragment. Instead, \( \text{V}_H \) and \( \text{V}_L \) domains from one scFv molecule associate with \( \text{V}_H \) and \( \text{V}_L \) domains from a second scFv molecule to form a bivalent dimer, termed a diabody (Hollinger et al., 1993). When the linker is shortened to <3 residues or when \( \text{V}_H \) and \( \text{V}_L \) domains are directly ligated to each other, scFv molecules associate to form a trimer, termed a triabody (Iliades et al., 1997; Kortt et al., 1997).

ScFvs of both \( \text{V}_H–\text{V}_L \)-linker–\( \text{V}_L \) and \( \text{V}_L–\text{V}_H \)-linker–\( \text{V}_H \) orientation have been produced by various research groups (Huston et al., 1991). Our laboratory has to date almost exclusively produced scFvs with the \( \text{V}_H \) domain at the amino terminus (Malby et al., 1993; Lilley et al., 1994; Coia et al., 1997; Iliades et al., 1997). In particular, NC10 anti-neuraminidase scFv antibody, in a \( \text{V}_H–\text{V}_L \)-linker–\( \text{V}_H \) orientation, has been used extensively to elucidate the oligomeric nature of scFvs with reduced linker lengths (Kortt et al., 1997; Atwell et al., 1999). To investigate the oligomerization phenomenon further, the NC10 scFv fragment was assembled in a reverse, \( \text{V}_H–\text{V}_L \)-linker–\( \text{V}_H \) orientation. The observed result showed that as the linker length between \( \text{V}_L \) and \( \text{V}_H \) was reduced, different patterns of oligomerization were observed compared with the original \( \text{V}_H–\text{V}_L \)-linker–\( \text{V}_H \) orientation. Furthermore, the direct ligation of \( \text{V}_L \) to \( \text{V}_H \) led to the formation of a tetramer, rather than to a trimer as observed for direct ligation of \( \text{V}_H \) to \( \text{V}_L \).

Materials and methods

Sequence numbering

Antibody residues were numbered according to Kabat et al. (1991) and for NC10 correspond exactly to Malby et al. (1993). Residues in the \( \text{V}_L \) domain of the scFv were superscripted with \( \text{L} \) and the residue number; for example, \( \text{Arg}^{107} \) signifies arginine in position 107 of the \( \text{V}_L \) domain. Similarly residues in the \( \text{V}_H \) domain of the scFv were superscripted \( \text{H} \) and the residue number.

General cloning procedures

Unless stated otherwise, all DNA manipulations were carried out according to standard protocols (Sambrook et al., 1989)
with reagents purchased from New England Biolabs. All polymerase chain reactions (PCRs) were performed with Pfu DNA polymerase (Stratagene). The PCR-amplified DNA fragments were digested with appropriate restriction enzymes, run on 1% (w/v) agarose gel and purified from the gel using BRESAclean purification kit (BresaTech). The purified DNA fragments were ligated into similarly prepared expression vectors using reagents and protocols supplied by Gibco-BRL. The ligation mixtures were transformed by the electroporation method (Dower et al., 1988) into Escherichia coli XL-1 Blue MRF’ cells (Stratagene) and recombinant clones were identified by colony PCR using primers complementary to 5’ and 3’ ends of recombinant gene inserts. All DNA sequences of various scFv constructs were verified using Dye Terminator Cycle Sequencing kits with AmpliTaq (PE Applied Bio-systems).

Construction of NC10 scFv gene fragments with 0 and 15 residue linkers
The previously described pPOW–NC10 scFv-15 (VH–VL) gene construct (Malby et al., 1993) was used as a source of VH and VL gene fragments. To generate the NC10 scFv-0 (VL–VH) synthetic gene, the VL and VH gene fragments were PCR-amplified using primers N4311 and N4341 for VL, and N4342 and N4293 for VH (Table I). The resulting VL and VH PCR products were gel-purified and joined into an scFv-0 using PCR overlap extension (Horton et al., 1989). To create the NC10 scFv-15 (VL–VH) synthetic gene, VL and VH gene fragments were PCR-amplified using primers N4311 and N4535 for VL, and N4534 and N4293 for VH (Table I). The resulting VL and VH gene fragments each contained part of the linker sequence at the 5’ and 3’ ends, respectively, as well as the BamHI restriction site which allowed for correct in-frame ligation of these V fragments via a linker sequence encoding 15 amino acid (GGGGS)5. The NC10 scFv-0 and scFv-15 (VL–VH) synthetic genes were then digested with NcoI and NotI restriction enzymes and cloned separately into a likewise digested pGC-4C2 E.coli vector (Coia et al., 1997) to create pGC-NC10 scFv-0 and scFv-15 plasmids (Figure 1a-i). This pGC-4C2 vector backbone incorporates two, rather than one, C-terminal FLAG peptide epitopes (Hopp et al., 1988) for improved purification efficacy of FLAG-fusion proteins. The resulting NC10 scFv-0 and scFv-15 gene constructs encoded N-terminal cleavable pelB periplasmic targeting signal followed by NC10 scFv and two C-terminal FLAG peptide epitope tags which were linked to VH domain at the C-terminus by three alanine residues and separated from one another by two alanine residues (Figure 1a-iii). The NC10 scFv-15 and scFv-0 (VL–VH) gene fragments were also inserted into a heat-inducible pPOW vector (Power et al., 1992). This was done by digesting the above described pGC-NC10 scFv-15 and scFv-0 (VL–VH) plasmids with NcoI and EcoRI restriction enzymes and ligating the scFv synthetic genes into a likewise digested pPOW vector. Similarly as for pGC constructs, the resulting gene fragment (Figure 1a-ii) was designed to express pelB signal sequence followed by NC10 scFv and two C-terminal FLAG tags.

Construction and cloning of NC10 scFv gene fragments with shorter linkers
The pGC-NC10 scFv-0 (VL–VH) gene construct (Figure 1a) was used for insertion of linkers of increasing length. This vector was digested simultaneously with XhoI and PstI restriction enzymes and the remaining plasmid DNA was purified using agarose gel. Five sets of synthetic oligonucleotides (Table I) were phosphorylated at the 5’ termini by incubation at 37°C in 30 min with 0.5 units of T4 polynucleotide kinase and 1 mM rATP (Biotech International) in 1× PNK buffer. Pairs of complementary phosphorylated oligonucleotides were pre-mixed in equimolar ratios to form DNA duplexes that encoded linkers of increased length. The one-residue linker construct was designed to encode a Ser residue in between codons for C-terminal VL ArgL107 and N-terminal VH GlhH11, whereas the two-, three-, four- and five-residue linker constructs were designed to encode additional glycine residues immediately preceding the Ser linker residue (Figure 1b). All five of these duplexes were designed to contain a ‘sticky end’ overlap compatible with XhoI and PstI restriction enzyme sites at the 5’ and 3’ ends, respectively. This allowed for direct cloning of these duplexes into XhoIPstI restricted pGC-NC10 scFv-0 (VL/VH) plasmid.

Fig. 1. (a) Schematic diagram of the NC10 scFv (VL/VH) expression unit in pGC (i) and pPOW (ii) vectors. (iii) Important parts of nucleotide and amino acid sequences of the NC10 scFv (VL–VH) unit in pGC and pPOW, outlining restriction sites for cloning (underlined), C-terminal region of pelB expression unit, N- and C-terminal sequence of VH gene (in bold), and two C-terminal octapeptide FLAG epitopes (in bold). (b) Amino acid sequences of the C-terminus of the VH domain, N-terminus of the VH domain and of the linker peptide (in bold) used in each of the NC10 scFv constructs.
Expression and purification of the scFvs

NC10 scFv-15 and scFv-0 gene constructs (V\textsubscript{L}–V\textsubscript{H}) were initially expressed in a heat-inducible pPOW vector (Power et al., 1992). Each pPOW-NC10 scFv construct was expressed in 1 l of 2YT/amp\textsubscript{100} as described by Malby et al. (1993), using E.coli strain TOP 10F\textsubscript{H11032} as host. The scFv-15 and scFv-0 proteins (V\textsubscript{L}–V\textsubscript{H}) were isolated from the periplasm as soluble fractions (supernatant) using an M2 anti-FLAG IgG-Sepharose column as described in the previous section. Eluted scFv proteins were dialysed against PBS and insoluble material removed by centrifugation. Recombinant scFv proteins were purified from soluble fractions (supernatant) using an M2 anti-FLAG IgG-Sepharose column (5×1 cm; Brizzard et al., 1994). The affinity column was equilibrated in PBS and bound proteins were eluted with 100 mM glycine, pH 3.0. The eluted proteins were neutralized with 1/10 volume of 1 M Tris–HCl (pH 8.0) and dialysed extensively against PBS–0.02% (w/v) sodium azide. Proteins samples were concentrated by ultrafiltration (Amicon) over a 10 kDa cut-off membrane (YM10, Diaflo) to ~1 mg/ml and stored at 4°C.

Expression and purification of NC10 scFvs (V\textsubscript{L}–V\textsubscript{H}) with shorter linkers

Each pGC-NC10 scFv (V\textsubscript{L}–V\textsubscript{H}) construct was expressed in 500 ml of 2YT/amp\textsubscript{100} + 0.1% d-glucose as described in Dolezal et al. (1995), using E.coli strain TOP 10F\textsuperscript{2} cells as host. Expression experiments were terminated 3 h post-induction and proteins were isolated from the periplasmic space using a modified method of Minsky et al. (1986). Briefly, cells were centrifuged at 5000 g for 10 min and supernatant was discarded. After washing the cell pellet in ice-cold PBS, cells were resuspended in 20 ml of ice-cold spheroplast buffer (100 mM Tris–HCl, pH 8.0, 0.5 M sucrose, 2 mM EDTA, 100 µg/ml lysozyme). After incubating the cell mixture on ice for 20 min, ice-cold half-strength spheroplast buffer (50 mM Tris–HCl, pH 8.0, 1 mM EDTA, 0.25 M sucrose) was added to a final volume of 100 ml. Cells were mixed gently and left on ice for a further 30 min. Periplasmic proteins were harvested by centrifugation at 10,000 g. The supernatant (100 ml) was sonicated briefly to shear the remaining DNA/RNA and filtered through a 0.45 µm filter. An Amicon concentrator unit (with YM10 membrane) was used to concentrate the periplasmic fraction to ~15 ml. Recombinant NC10 scFvs (V\textsubscript{L}–V\textsubscript{H}) proteins were affinity purified from the periplasmic fraction using an M2 anti-FLAG IgG-Sepharose column as described in the previous section. Eluted scFv proteins were dialysed against PBS–0.02% sodium azide, concentrated by ultrafiltration (YM10, Diaflo) to ~1 mg/ml and stored at 4°C.

Biochemical and biological characterization of NC10 scFv proteins

The purity of the NC10 scFvs was monitored by SDS–PAGE and Western blot analysis as described previously (Kortt et al., 1994). The concentrations of the scFv fragments were determined spectrophotometrically, using values for the extinction coefficient ($\varepsilon$) at 280 nm of 1.66 for V\textsubscript{L}–V\textsubscript{H} scFvs and 1.70 for V\textsubscript{H}–V\textsubscript{L} scFvs calculated as described by Gill and von Hippel (1989). The relative molecular mass of each affinity-purified NC10 scFv was estimated using size-exclusion chromatographic columns (Superose 12 HR10/30 and/or Superdex 200 HR 10/30, Pharmacia) on an HPLC system (Bio-Rad Model 700) at 21°C in PBS which were previously calibrated with Bio-Rad Gel Filtration Standard proteins. The flow-rate was 0.5 ml/min, and the absorbance of the effluent stream was monitored at both 214 and 280 nm. Elution times of various V\textsubscript{L}–V\textsubscript{H} scFv oligomers were compared with those already established for V\textsubscript{H}–V\textsubscript{L} scFv monomers, dimers and trimers (Kortt et al., 1994, 1997; Atwell et al., 1999).
Formation of complexes with 3–2G12 anti idiotypic Fab

Fab' fragments of the NC10 anti-idiotypic antibody, 3–2G12, were prepared as described in Kortt et al. (1997). Purified NC10 scFv-0, scFv-1, scFv-2 and scFv-5 (V_L–V_H) proteins were mixed with a small molar excess of 3–2G12 anti-idiotypic Fab', as described previously by Kortt et al. (1997). The complexes were separated from excess Fab' by size-exclusion chromatography on Superdex 200 (HR 10/30) in PBS, pH 7.4, with a flow-rate of 0.5 ml/min. The column had previously been calibrated with uncomplexed scFvs and 3–2G12 Fab'.

Electron microscopy

Complexes of scFv-0 and scFv-1, scFv-2 and scFv-5 with 3–2G12 Fab' and the complex of influenza virus neuraminidase (soluble tetrameric extracellular domain) with NC10 Fab' (Malby et al., 1994) were examined by electron microscopy (EM). EM imaging and data analysis were performed as described previously (Lawrence et al., 1998; Atwell et al., 1999).

Molecular modeling

Computer-generated models of NC10 scFv triabodies and tetrabodies were constructed using Fv modules that corresponded to the coordinates of the NC10 Fv domain in PDB entry 1NMB (Malby et al., 1994, 1998). Fv modules were manipulated as rigid bodies with the O molecular graphics package (Jones et al., 1991). Triabody structure corresponded to the model described by Kortt et al. (1997) comprising three Fv modules with threefold symmetry. Tetrabodies comprised four Fv modules with fourfold symmetry. No attempt was made to model conformational changes in the Fv domains.

Results and discussion

Expression of NC10 scFv-15 and scFv-0 constructs

For direct comparison with the original NC10 scFv-15 and scFv-0 (V_H–V_L) constructs (Malby et al., 1993; Kortt et al., 1997), the NC10 scFv-15 and scFv-0 (V_L–V_H) gene fragments were initially constructed and expressed in a heat-inducible expression vector pPOW (Figure 1a–iii). As for V_H–V_L orientation, in the V_L–V_H construct of the NC10 scFv, the V_L domain was linked to the V_H domain using the classical linker design of Huston et al. (1988, 1991), whereby the codons for C-terminal V_L Arg_107 were linked to the codon for N-terminal V_H Gln_1 via a 15 amino acid residue linker ([G4S]3; Figure 1b). The Arg_107 was defined from the NC10 scFv crystal structure to be the last residue in the V_L that made intra-domain contacts within V_L and additional residues therefore formed a true linker (Malby et al., 1998). In case of scFv-0, the C-terminal V_L Arg_107 was ligated directly to V_H Gln_1 (Figure 1b).

The pPOW/NC10 scFv-15 and scFv-0 (V_L–V_H) constructs were expressed under the same conditions as those used previously for the V_H–V_L constructs (Malby et al., 1993; Kortt et al., 1997). As was the case for the V_H–V_L proteins, the majority of expressed scFv-15 and scFv-0 (V_L–V_H) proteins were located in the periplasm as insoluble protein aggregates which were solubilized by extraction with 6 M GuHCl and on dialysis refolded into soluble, functional scFv entities. Furthermore, there was no apparent difference in expression levels between the V_L–V_H and V_H–V_L scFvs proteins. The resulting soluble scFvs were purified by affinity chromatography on an M2 anti-FLAG IgG-Sepharose column. Yields of 2–5 and 1–2 mg of soluble affinity-purified protein per litre of shake-flask culture were typically obtained for scFv-15 and scFv-0, respectively.

Biophysical analysis of NC10 scFv-15 and scFv-0 proteins

Samples of affinity-purified scFv-15 proteins (V_L–V_H and V_H–V_L) were shown by SDS–PAGE to comprise essentially homogeneous scFv preparations with a main protein band at ~27.0 kDa for V_H–V_L and ~28.5 kDa for V_L–V_H (data not shown). This apparent difference in protein mobility on SDS–PAGE was attributed to the additional sequence associated with the second FLAG epitope that has been added to the C-terminus of scFv V_H–V_L proteins (see Materials and methods and Figure 1a). Size-exclusion chromatography on a calibrated Superose 12 column showed that the oligomeric status of the NC10 V_H–V_L and V_L–V_H scFv-15 proteins was similar in solution (Figure 2). Gel filtration of affinity-purified scFv-15 V_L–V_H protein showed the presence of two main peaks with apparent molecular masses of ~28.5 kDa (major peak) and ~57.0 kDa (minor peak), corresponding to scFv-15 monomer and dimer, respectively. The scFv-15 V_H–V_L protein when analysed on the same column eluted mainly as monomer at ~27.0 kDa with a small amount of dimer at ~54.0 kDa. The relative differences in elution times for V_H–V_L and V_L–V_H monomers and dimers were again attributed to the additional FLAG sequence at the C-terminus of scFv-15 V_L–V_H protein. Both scFv-15 protein samples also contained traces of higher order oligomers, but because of small quantities of these species it was not possible to assign their oligomeric status unequivocally. Similarly as shown for NC10 scFv-15 V_H–V_L protein (Kortt et al., 1994), the formation of NC10 scFv-15 V_L–V_H dimers and higher molecular mass multimers was induced by storing the sample at higher concentrations (>1 mg/ml) or by repeated freezing and thawing (data not shown).

In contrast to scFv-15 proteins, size-exclusion chromatography of scFv-0 (V_L–V_H and V_H–V_L) proteins demonstrated a significant difference in elution profiles (Figure 3a). The chromatography of the NC10 scFv-0 (V_L–V_H) protein yielded a major protein peak eluting at ~21.9 min with a distinct...
Expression of NC10 V\textsubscript{L}–V\textsubscript{H} scFvs with variable linker length

To investigate further the effect of linker length upon multimerization of NC10 scFvs in reverse (V\textsubscript{L}–V\textsubscript{H}) orientation, seven different NC10 scFv constructs with linkers from 15 to 0 amino acid residues were assembled in pGC secretion vector (Coia et al., 1997) as described in Materials and methods. This vector was chosen for construction and expression of these scFvs with shortened linkers in preference to pPOW vector because of its capacity to produce soluble and active scFv proteins in the bacterial periplasm. A schematic diagram depicting the general outline of these short-linked NC10 scFv expression units in pGC is shown in Figure 1b. The pGC NC10 scFv V\textsubscript{L}–V\textsubscript{H} constructs were expressed in E.coli TOP 10F′ cells using the expression protocol of Dolezal et al. (1995). The expressed scFv products were isolated from the periplasmic fraction by affinity chromatography as described in Materials and methods. The yield of soluble affinity-purified protein decreased progressively as the linker length was shortened, from 2 mg/l bacterial culture for scFv-15 to 0.5 mg/l bacterial culture for scFv-0.

**Molecular mass analysis of scFv proteins with shorter linkers**

SDS–PAGE analysis of affinity-purified NC10 scFv-0, scFv-1, scFv-2, scFv-3, scFv-4, scFv-5 and scFv-15 (V\textsubscript{L}–V\textsubscript{H}) protein samples showed that the scFvs comprised predominantly a single component of apparent molecular mass ~27.5–28.5 kDa, as expected for this series of proteins (Figure 4). When these scFvs were subjected to analysis by size-exclusion chromatography on a calibrated Superdex 200 column, a number of major protein peaks were observed with a significant variation in elution times consistent with the presence of scFv oligomers (Figure 5). Elution profiles for scFv-5, scFv-4 and scFv-3 (Figure 5a) showed the presence of a single major peak with an elution time consistent with a molecular mass of ~57 kDa expected for a dimer. This peak eluted at the same time as the dimer peak in scFv-15 (V\textsubscript{L}–V\textsubscript{H}) preparations (Figure 5b). The minor higher molecular mass species observed in these three profiles (Figure 5a) eluted at the same elution times as the zero-linked NC10 V\textsubscript{L}–V\textsubscript{H} trimer and tetramer peaks (Figure 5b). These findings are consistent with those observed for the NC10 scFv-5, scFv-4 and scFv-3 (V\textsubscript{L}–V\textsubscript{H}) proteins which mainly formed dimer and a small amount of trimer (Atwell et al., 1999). However, no tetramer species was observed for scFv-5, -4 and -3 (V\textsubscript{H}–V\textsubscript{L}) proteins. Interestingly, in the case of scFv-5 (V\textsubscript{H}–V\textsubscript{L}) protein, consistently larger yields of tetramer species were observed and this allowed for the purification and subsequent characterization of this tetramer by gel filtration on a Superdex 200 column and by electron microscopy (see below). This scFv-5 tetramer was, however, relatively unstable as it partially reverted back to dimer (40% dimer after 24 h at 4°C). Similarly, a small amount of purified scFv-5 dimer species converted to tetramer (~5–10% after 24 h at 4°C).

---

**Fig. 3.** Size-exclusion chromatography on a calibrated Superdex 200 HR 10/30 column of affinity-purified NC10 scFv-0 proteins. Columns were equilibrated in PBS, pH 7.4, and the flow-rate was 0.5 ml/min. (a) Shows the scFv-0 V\textsubscript{L}–V\textsubscript{H} tetramer (solid line) eluting at 21.9 min with a trimer shoulder on the trailing edge. Superimposed on the same scale is a run for scFv-0 V\textsubscript{H}–V\textsubscript{L} trimer (dashed line) eluting at 24.1 min. (b) Shows separation of scFv-0 tetramer from scFv-0 trimer using two Superdex 200 HR 10/30 columns linked in tandem.

---

**Fig. 4.** Reducing SDS–PAGE of affinity-purified NC10 scFvs (V\textsubscript{L}–V\textsubscript{H}) stained with Coomassie Brilliant Blue G-250. Lanes: 1, scFv-0; 2, scFv-1; 3, scFv-2; 4, scFv-3; 5, scFv-4; 6, scFv-5; 7, scFv-15.
Size-exclusion chromatographic profiles of scFv-1 and scFv-2 yielded a series of peaks consistent with an equilibrium mixture of scFv dimers, trimers, tetramers and higher molecular mass multimers (Figure 5c). In the case of scFv-2 this equilibrium mixture was biased towards the smaller scFv entities whereas in case of scFv-1 the equilibrium mixture shifted towards higher molecular mass multimers. Analysis of entire scFv-1 and scFv-2 protein mixtures in 60% ethylene glycol and at lower concentration (<100 µg/ml) showed that the higher molecular mass multimers dissociated primarily into dimers and trimers (data not shown). This suggests that the dimers and trimers of scFv-2 and particularly scFv-1 are relatively unstable and as a result prefer to assemble into higher molecular mass multimers as the linker length is reduced from two to one amino acid residue.

As described previously (Figure 1), when compared with the original NC10 scFv V<sub>L</sub>–V<sub>H</sub> proteins, a second FLAG sequence was included in our constructs to improve the purification efficiency of various NC10 V<sub>L</sub>–V<sub>H</sub> proteins. To confirm that this additional FLAG sequence does not affect the state of oligomerization, several pGC-NC10 V<sub>L</sub>–V<sub>H</sub> clones containing a single FLAG sequence were also constructed and expressed in E.coli. The ‘single-FLAG’ NC10 V<sub>L</sub>–V<sub>H</sub> proteins with no, one, two and 15 residue linkers were purified on an anti-FLAG affinity column and analysed by gel filtration chromatography. The resulting gel filtration profiles indicated no apparent difference in oligomerization properties between ‘single-FLAG’ and ‘double-FLAG’ NC10 V<sub>L</sub>–V<sub>H</sub> proteins (data not shown). The presence of minor bands at ~26 kDa in affinity-purified NC10 V<sub>L</sub>–V<sub>H</sub> samples (Figure 4) provided further evidence that the additional FLAG peptides do not affect the oligomeric status of these proteins. These bands occur owing to a gradual cleavage of FLAG peptides from the C-terminus of NC10 scFv proteins during extended storage at 4°C (demonstrated by Western blot analysis with the anti-FLAG M2 antibody, data not shown). The presence of these bands, however, had no significant effect on the gel filtration profiles, and hence multimeric forms, of these NC10 scFvs (data not shown).

**Formation of complexes with 3–2G12 anti-idiotype Fab’ fragments**

To analyse further the multimeric status of NC10 scFvs (V<sub>L</sub>–V<sub>H</sub>), the Fab’ fragment of 3–2G12 anti-idiotype monoclonal antibody, which competes with influenza virus neuraminidase for binding to NC10, was used to form complexes with various scFv V<sub>L</sub>–V<sub>H</sub> proteins. Affinity-purified scFv-5 dimer and scFv-5 tetramer, scFv-0 tetramer as well as the scFv-2 and scFv-1 mixtures were complexed with 3–2G12 Fab’ fragment and the resulting complexes analysed by gel filtration (data not shown) and by electron microscopy (see below). In the case of scFv-5, the tetramer species was separated from dimer by gel filtration just prior to binding to 3–2G12 Fab’. Complexes were formed by mixing dimers (for scFv-5) and tetramers (for scFv-5 and scFv-0) in a 1:2 and 1:4 molar ratio, respectively, with the 3–2G12 Fab’ fragment. The Fab’ was kept in slight excess to these ratios to ensure complete decoration of all antigen binding sites present. The resulting scFv dimer–Fab’ and scFv tetramer–Fab’ complexes were then purified by size-exclusion chromatography on a Superdex 200 column (data not shown). There was no evidence of unbound scFv tetramer (for scFv-0 and scFv-5) or scFv dimer (for scFv-5) in any of these preparations. The scFv-5 dimer–Fab’ complex eluted at an elution time corresponding to a molecular mass of ~160 kDa, which is consistent with the mass of a complex of two Fab’ molecules (~52 kDa each) and one scFv-5 dimer (~55 kDa). The scFv tetramer–Fab’ complex (both scFv-0 and scFv-5) eluted at an elution time corresponding to a molecular mass of ~320 kDa, which is consistent with the mass of a complex of four Fab’ molecules and one scFv tetramer (~108 kDa). The estimated molecular mass of scFv–anti-idiotype Fab’ complexes is consistent with the prediction that scFv-5 dimers are bivalent and bind two Fab’ fragments and that scFv-0 (and scFv-5) tetramers are tetravalent and bind four Fab’ fragments. Minor peaks were
sometimes observed during purification of these scFv-5 dimer–Fab’ and scFv-5 tetramer–Fab’ complexes. These ‘contamination’s’ occurred as a result of the rapid equilibrium between scFv-5 tetramer and dimer species. Thus, during purification of scFv-5 tetramer–Fab’ complex a small but significant amount of scFv-5 dimer–Fab’ complex was detected. Similarly, trace quantities of scFv-5 tetramer–Fab’ complex were present in scFv-5 dimer–Fab’ complex. These ‘contaminations’ were consistently present in various scFv-5 dimer–Fab’ and tetramer–Fab’ complex preparations in spite of the fact that pure samples of scFv-5 dimer and tetramer proteins were purified by gel filtration immediately prior to their binding to antidiotypic Fab’. Similarly, during purification of the scFv-0 tetramer–Fab’ complex, a distinct shoulder was observed on the trailing edge of the peak. This indicated the presence of scFv-0 trimer–Fab’ complex that resulted from rapid interconversion of the gel filtration-purified scFv-0 tetramer into a mixture of tetramer and trimer. Whilst size-exclusion chromatography on Superdex 200 separated the scFv-5 dimer–Fab’ complex from scFv-5 tetramer–Fab’ complex, it was not possible to separate thoroughly the scFv-0 trimer–Fab’ complex from scFv-0 tetramer–Fab’ complex on this column. Consequently, the scFv-0 trimer–Fab’ complex was always present in preparations of the scFv-0 tetramer–Fab’ complex. The gel filtration-purified scFv–Fab’ complexes were stable and could be stored either at 4°C for several weeks or frozen and thawed as required.

Electron microscopy of NC10 scFvs (V_L–V_H) complexed with 3–2G12 Fab’ fragments

The size-exclusion-purified scFv dimer–3–2G12 Fab’ and scFv tetramer–3–2G12 Fab’ complexes were imaged by electron microscopy. Images of the scFv-5 dimer–Fab’ complex appeared predominantly as boomerang-shaped projections (Figure 6a). These images were identical with those observed previously for NC10 scFv-5 (V_H–V_L) diabodies complexed with 3–2G12 Fab’, which showed projections consistent with two Fab’ molecules extending outwards from the antigen binding sites (Lawrence et al., 1998; Atwell et al., 1999). Despite the potential ‘elbow’ flexibility between Fv and C modules in the Fab’, each Fab’ arm appeared as a relatively rigid, linear molecular rod (Tulloch et al., 1986; Lawrence et al., 1998). Images of complexes formed by association of scFv tetramer and 3–2G12 Fab’ (Figure 6b and c) appeared as X-shaped projections which were similar to complexes of NC10 Fab’ and influenza neuraminidase tetramers formed under the same imaging conditions (Figure 6d). As discussed previously, the scFv-0 tetramer–Fab’ preparations were partially contaminated with the scFv-0 trimer–Fab’ complex. As a result, some of the images shown in Figure 6c appeared as Y-shaped and V-shaped projections identical with those observed for scFv-0, -1 and -2 (V_H–V_L) triabody–Fab’ complex imaged previously (Lawrence et al., 1998; Atwell et al., 1999). The Y-shaped projections were previously interpreted as tripod-shaped objects (viewed from above), with the distal ends of
the three Fab' molecules in contact with the carbon film (Lawrence et al., 1998). The V-shaped projections, on the other hand, were interpreted as tripod-shaped objects lying on their sides on the carbon film, with one Fab' molecule extending upward and partially out of the stain (Lawrence et al., 1998). Based on these interpretations, similar reasoning can be applied to the interpretation of tetramer complexes. Thus, for example, X-shaped objects lying on their sides on the carbon film may appear as tripods with one Fab' leg extending upward and partially out of the stain. These interpretations, however, have to be treated cautiously owing to the presence of contaminating scFv-0 trimer complexes. Consequently, some of these Y-shaped and V-shaped projections may, in fact, be 'true' trimers which may ultimately interfere with the correct interpretation of electron micrographs of scFv–Fab' tetramer complexes. Owing to this trimer contamination and more importantly owing to increased complexity of scFv–Fab' tetramer complexes (in comparison with trimers and dimers), further analysis of these EM images, including accurate measurement of take-off angles for Fab' arms, would be highly speculative and consequently were not performed.

In the case of affinity-purified scFv-1 and scFv-2 protein preparations that consisted of mixtures of dimers, trimers, tetramers and higher molecular-mass multimers (Figure 5c), a large excess of 3–2G12 Fab' fragment was added to ensure complete saturation of all potential binding sites. The resulting scFv–Fab' protein complexes (both scFv-1 and 2) were separated from unbound excess Fab' by size-exclusion chromatography on a Superdex 200 column (data not shown). There was no evidence of any unbound scFv tetramer, trimer and dimer in either of these two preparations. Owing to the lack of resolution between higher molecular mass scFv multimers and the smaller scFv dimer–Fab' and scFv trimer–Fab' complexes, individual complexes were not isolated. Consequently, the scFv-1 and scFv-2 complexes were collected as a mixture of high molecular mass peaks and analysed directly by electron microscopy (Figure 6e and f). The resulting EM images revealed an array of different projections that included boomerang-shaped dimers, Y-shaped trimers, X-shaped tetramers and 'star'-shaped multimers binding a varying number of Fab' arms. The frequency distribution of these images was consistent with the relative proportional distribution of scFv oligomeric components observed on size-exclusion chromatography (Figure 5c). Thus, a large number of 'star'-shaped projections was observed for scFv-1–Fab' complex images in which higher molecular mass multimers were the favoured species (Figure 6e). This contrasted with the scFv-2–Fab' complex images in which boomerang-shaped dimers were the predominant species (Figure 6f). These images clearly demonstrated that the transition between NC10 V_L–V_H dimers, trimers and tetramers was not as distinct as for their V_H–V_L isomers which showed a precise transition between dimers and trimers when the linker length was reduced from three to two residues (Atwell et al., 1999).

**Molecular models**

To understand the relative size and orientation of EM images, molecular models of an NC10 scFv-0 (V_L–V_H) triabody and tetrabody were constructed with three- and fourfold symmetry (Figure 7a and b), respectively, and compared with the known structure of a neuraminidase tetramer (Figure 7c). Both models shown in Figure 7a and b are represented by a single rigid body structure each showing one plausible conformation for triabody and tetrabody, respectively. Neither of these two models contain C-terminal FLAG sequences and neither has been energy minimized or analysed extensively by molecular dynamics. These models are included to provide a single molecular view that represents the Fv module sizes and relative orientations. The trimer model is based on the previously described model of Kortt et al. (1997) and assumes that three symmetric Fv heads with polypeptide tails are assembled in a cyclic head-to-tail fashion. The scFv tetrabody model has a similar domain arrangement with a symmetric tetrameric
conformation and fewer steric constraints than the triabody. Although the neuraminidase tetramer shows a similar overall size to tetrabodies, the subunits are closely packed in a stable and rigid association. In contrast, many different associations between Fv modules are possible in both triabodies and tetrabodies, but particularly in tetrabodies. The flexibility is best observed in the EM images of tetrabodies complexed with anti-idiotype Fab' (Figure 6b and c) compared with neuraminidase complexed with Fab' (Figure 6d). The neuraminidase complexes appear as rigid X-shaped projections with a fixed take-off angle of the four Fab’ arms extending radially out of the page from the core neuraminidase tetramer, consistent with the high-resolution crystal structure (Malby et al., 1994, 1998). In contrast, the X-shaped projections of tetrabody complexes show considerable flexibility in the take-off angles of anti-idiotype Fab’ arms, presumably owing to flexible orientations of the Fv modules. The central regions in these X-shaped EM images frequently have an increase in stain density in this area. This observation is consistent with four symmetrical subunits associating together to form a central cavity (viewed from above). Comparison of molecular models of NC10 scFv-0 V_{L}–V_{H} and V_{H}–V_{L} trimers suggests that steric restrictions may cause the preference for four instead of three scFv molecules for NC10 scFv-0 V_{L}–V_{H}. In particular, loop L55–L61 (especially residues L55–L58) appears to clash with residues L15–L17 in V_{L}–V_{H} triabodies (Figure 7a), whereas the structurally similar loop (H58–H67) in V_{H}–V_{L} tetrabodies does not (Atwell et al., 1999).

**ScFv multimers**

The oligomerization of scFvs has been described as a manifestation of the domain swapping phenomenon (Arndt et al., 1998) where one domain of a monomeric protein is being replaced by the same domain from an identical protein chain (Bennett et al., 1994, 1995). The extent of such oligomerization is driven by various factors including expression conditions, folding methods, association constant for domain–domain interaction and steric accessibility of domain binding interfaces (Arndt et al., 1998). The domain swapping phenomenon is observed for the NC10 scFv-15 (V_{L}–V_{H}) examined in this study (Figure 5b) which consists of a mixture of monomer, dimer and small amounts of higher oligomers (mainly trimers and tetramers). Shortening of the linker length between V_{L} and V_{H} domains in NC10 scFv progressively prevents the formation of the smaller scFv entities, whilst still allowing formation of higher order oligomers. Thus, for example, the length of the linker in NC10 scFv-5, scFv-4 and scFv-3 prevents a monomeric arrangement but allows for the formation of dimers, trimers and tetramers (Figure 5a) that are present in the NC10 scFv-15 preparation in relatively low amounts (Figure 5b). Apart from our studies on the NC10 scFv, the exact linker length required for the diabody–triabody–tetrabody transition for other scFv molecules has been poorly defined. Le Gall et al. (1999) described an anti-CD19 (HD37) scFv in V_{H}–V_{L} orientation where the scFv-18 protein formed monomer, dimer and small amount of tetramer, the scFv-10 formed mainly dimer with some tetramer, the scFv-1 formed exclusively tetramers and the scFv-0 formed trimers. The authors, however, did not report on scFvs with three and two amino acid residues and consequently did not define the transition between a dimer and a tetramer. Interestingly, in the case of this anti-CD19 scFvs the formation of a tetramer with scFv-1 and a trimer with scFv-0 is an unexpected observation based on previous data (Atwell et al., 1999) and the present results. However, binding assays with anti-CD19 scFv-0 triabody suggested only a monovalent interaction with the cell surface anchored antigen (Le Gall et al., 1999), indicating that this triabody may not be correctly folded to allow formation of a trivalent molecule with three binding sites.

It has been documented that V_{H} and V_{L} domains are not related by a rotation of 180° about the pseudo twofold axis (parallel to the interface), but only by about 170° (Huston et al., 1991; Padlan, 1994). Therefore, the C-terminus of V_{L} is further away from the N-terminus of V_{H} (~39–43 Å) than the C terminus of V_{H} is from the N-terminus of V_{L} (~32–34 Å). Consequently, when the two isomers are linked with identical linkers the greater constraints in the V_{L} to V_{H} construct would be expected to favour a larger proportion of higher order oligomers (Plückthun and Pack, 1997). This expectation is contradicted in the present study which shows that the scFv-2 and scFv-1 V_{L}–V_{H} protein preparations contain functional dimers whereas their V_{H}–V_{L} isomers are trimeric. It is tempting to speculate that in the V_{L}–V_{H} proteins the C-terminal residues of V_{L} domain and/or N-terminal residues of V_{H} domain provide an additional flexibility in the linker region, which drives this transitional formation of dimers, trimers and tetramers. However, the association of Fv modules is unlikely to depend on the linker length and linker flexibility alone. The current evidence from crystal structures of scFv dimers and trimers in the V_{H}–V_{L} orientation suggests that interface interaction across Fv modules in diabodies and triabodies is almost exclusively between V_{H} domains (Perisic et al., 1994; Hollinger et al., 1996; Pei et al., 1997; Malby et al., 1998). The detailed molecular interactions between V-domain interfaces, Fv interfaces and the terminal residues held in contact within the V-domain framework are obviously unique for each antibody and therefore will affect both flexibility and stoichiometry of diabodies, triabodies and tetrabodies.

**Conclusions**

The main aim of this study was to compare NC10 scFv proteins, assembled in the reverse V_{L} to V_{H} orientation, with the previously studied NC10 scFv V_{H}–V_{L} proteins (Malby et al., 1993; Kortt et al., 1997; Atwell et al., 1999). Of special interest was the effect of linker length upon multimerization of these ‘reverse’ scFvs. Based upon size-exclusion chromatographic profiles and EM imaging of scFv and anti-idiotype Fab’ complexes, the following conclusions were reached: (1) in the V_{H}–V_{L} orientation, the NC10 scFv-15 V_{L}–V_{H} protein consists of a mixture of monomer and dimer; (2) dimers (diabodies) are the predominant conformation for scFv V_{L}–V_{H} proteins with five, four and three residue linkers; (3) unlike the scFv-0 V_{H}–V_{L} protein which forms a stable trimer (triabody), the scFv-0 V_{L}–V_{H} protein consists of a mixture of tetramer (major component) and trimer (minor component); (4) the transition between NC10 scFv V_{L}–V_{H} diabodies, triabodies and tetrabodies is not as distinct as for V_{H}–V_{L} orientation since reducing the linker length in V_{L}–V_{H} from three to two residues (or one residue) generates mixtures of dimers, trimers, tetramers and higher molecular mass multimers. The NC10 scFv-0 V_{L}–V_{H} molecule is the second report of a tetrameric scFv. It is reasonable to expect that the tetrabody conformations will be formed for many antibodies other than the two current reports for NC10 (here) and for HD37 (Le Gall et al., 1999). Given the demand for novel multivalent
antibody constructs in cancer therapy and other in vivo applications (Hudson, 1999), it is important to establish the generic rules for scFv association that promote high-level and cost-effective production. Consequently, studies investigating the reasons why the NC10 scFv-0 VL-VH molecule forms a tetramer rather than an expected trimer are in progress.

Acknowledgements
The authors thank Professors D.Metzger and R.Webster for providing the 3–2G12 and NC10 hybridoma cell lines, Dr D.Hewish for preparing the 3–2G12 antibodies, J.Burns for assistance with protein purification and chromatographic analysis and Drs N.McKern and S.Nuttall for critical reviews of the manuscript.

References