Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Challenging semi-bootstrapping molecular-replacement strategy reveals intriguing crystal packing of rhizavidin

Amit Meir and Oded Livnah*

Department of Biological Chemistry, The Alexander Silverman Institute of Life Sciences, The Wolfson Centre for Applied Structural Biology, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel

Correspondence e-mail: oded.livnah@huji.ac.il

Received 11 January 2010
Accepted 2 February 2010

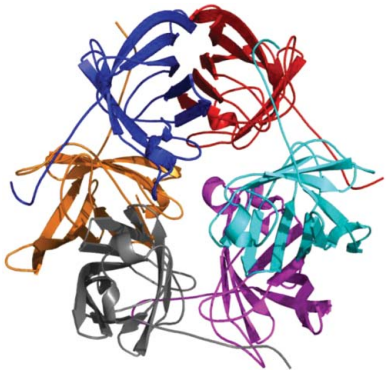
PDB References: rhizavidin, 3ew1; rhizavidin–biotin complex, 3ew2.

The structure of rhizavidin, the first dimeric member of the avidin family which maintains high affinity towards biotin, was determined to high resolution by SeMet SAD. Consequently, the structure of the rhizavidin–biotin complex was determined by molecular-replacement methods using the apo structure as the search model; this ran into complications and required combined programs as well as bootstrapping approaches. Although present as a dimer in solution, rhizavidin packs as unique oligomers in both crystal forms. The novel insights derived from the unique molecular-replacement procedure and the crystal-driven oligomeric forms in this work may have utililty in biotechological and nanotechnological applications.

1. Introduction

Rhizavidin from the bacterium *Rhizobium etli* is a recently discovered unique member of the avidin family. Avidins such as egg-white avidin and the bacterial streptavidin from *Streptomyces avidinii* are well characterized and are known for their extremely high affinity towards the vitamin biotin (K_d values of $\sim10^{-16}$ *M* for avidin and $\sim10^{-14}$ *M* for streptavidin; Green, 1975, 1990). All avidins exhibit similar tertiary structures (Fig. 1*a*) consisting of eight antiparallel β-strands forming a β-barrel, with the biotin-binding site at the wide end of the barrel. The differences between the avidins involve variations in the size, composition and conformation of the loops connecting the strands (Livnah *et al.*, 1993; Weber *et al.*, 1987; Hendrickson *et al.*, 1989). The quaternary structures consist of homo-tetrameric assemblies constructed as a dimer of dimers (Kurzban *et al.*, 1991). The three monomer–monomer interactions are defined as 1–2, 1–3 and 1–4 as designated by Livnah *et al.* (1993) (Fig. 1), where the latter represents the sandwich-like dimer with the highest monomer–monomer contact surface (average values of 1850 and 1535 Å^2 for avidin and streptavidin, respectively; Fig. 1*c*). In tetrameric avidins the 1–2 interaction is crucial since a critical tryptophan located in the L7,8 loop is donated to the biotin-binding site from an adjacent monomer and thus contributes to the high affinity and the tetrameric integrity (Fig. 1*b*; Laitinen *et al.*, 1999; Freitag *et al.*, 1998; Sano *et al.*, 1997). However, rhizavidin is currently the only member of the avidin family that has a homodimeric rather than a tetrameric structure while maintaining its high affinity towards biotin (Fig. 1*e*; Helppolainen *et al.*, 2007). In addition, rhizavidin lacks the critical Trp from the L7,8 loop that forms a lid for the biotin-binding site in tetrameric avidins.

The high affinity of the avidin–biotin system is being utilized in numerous biochemical, biotechnological and nanotechnological applications, including protein labelling, drug delivery and immuno-labelling (Wilchek & Bayer, 1990; Bayer & Wilchek, 1990; Paganelli *et al.*, 2001; Urbano *et al.*, 2007; Dehlinger *et al.*, 2007). The advantage of the four high-affinity biotin-binding sites could also become a drawback owing to undesired cross-linking, which could be resolved by high-affinity dimeric and/or monomeric avidins. However, when



© 2010 International Union of Crystallography
All rights reserved

tetrameric avidins are dissociated into their corresponding monomers the affinity towards biotin decreases by 7–8 orders of magnitude (Määttä *et al.*, 2008; Laitinen *et al.*, 2006, 2007). Such high-affinity monomeric and dimeric avidins could be achieved by rational design. Hence, alteration of the intermonomeric contact surfaces from a hydrophobic to a polar nature, followed by binding-site mutagenesis, could result in high affinities comparable to those of the tetrameric molecules. Another approach includes screening for new avidins from different species which may contain the desired properties of molecular assembly and affinity. In this regard, the high-affinity dimeric rhizavidin from *R. etli* provides a substantial leap forward in the field of avidin technology and protein design.

In order to interpret the components which permit rhizavidin to maintain its high affinity despite being a dimer (Helppolainen *et al.*, 2007), we have determined the crystal structures of rhizavidin in apo and biotin-complexed forms (Meir *et al.*, 2009). The apo structure was determined by SAD using an SeMet derivative. Subsequently, the biotin-complexed structure was determined *via* molecular-replacement methods using the apo structure as the search model, but this was not a simple task. The stable rhizavidin dimer is formed by the 1–4 monomer–monomer interaction, while the 1–3 and 1–2 interactions are not present. The consequent lack of the capping Trp residue results in partial availability of bound biotin to solvent, yet the high affinity is maintained (Meir *et al.*, 2009).

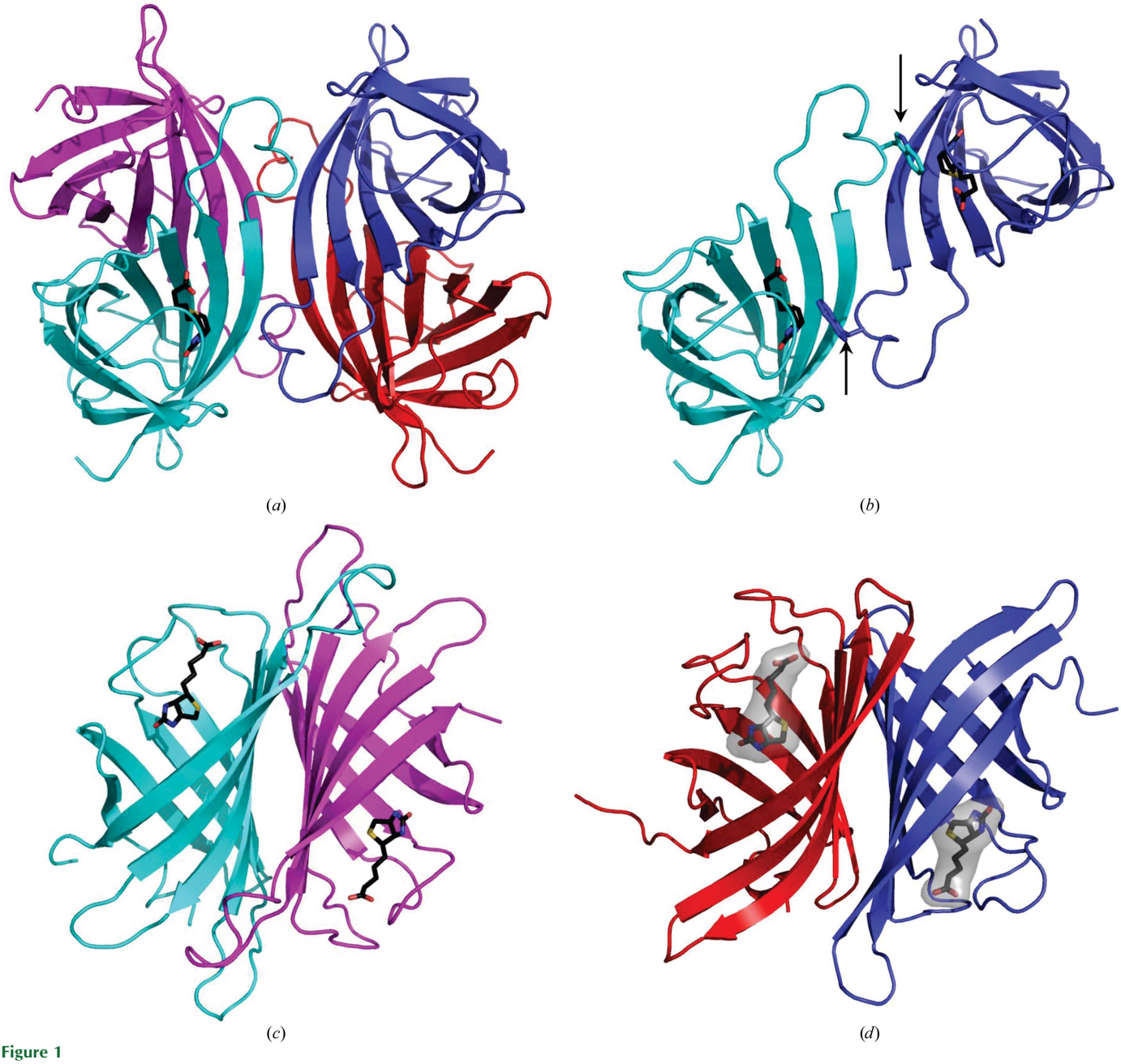


Figure 1
Ribbon presentation of the quaternary structures of tetrameric avidin and dimeric rhizavidin. (*a*) The tetrameric quaternary structure of avidins, with the monomers shown in red, blue, cyan and magenta. The biotin ligand is shown in black and for clarity occupies only one monomer (cyan). (*b*) The individual 1–2 interaction which is stabilized by Trp residues contributed from neighbouring monomers (indicated by short arrows). (*c*) The 1–4 sandwich-like intimate monomer–monomer interaction. (*d*) The rhizavidin dimer shown in red and blue with biotin molecules in the binding site. The interaction is highly similar to that of the 1–4 interaction in the tetrameric avidins. All molecular-graphics figures were generated using *PyMOL* (DeLano, 2002).

In this report, we describe the structure solution *via* molecular replacement of the rhizavidin–biotin complex, which was not straightforward and required combined methods as well as several search models and a bootstrapping approach. The difficulties in molecular-replacement solution arose owing to differences in crystal packing between the apo and biotin-complexed forms. Unlike the crystal packing of apo rhizavidin, in which the asymmetric unit contains a bagel-shaped arrangement of three dimers (six monomers), the asymmetric unit of the biotin complex contains seven rhizavidin monomers that pack in a different manner. Our results provide an analysis of the intriguing crystal packing of the two rhizavidin forms: the aesthetic and symmetrical configuration of the apo rhizavidin in its hexameric form and the unique heptameric symmetry-dependent form of the biotin complex. These crystallization-induced molecular configurations could be exploited in biotechnological or nanotechnological applications.

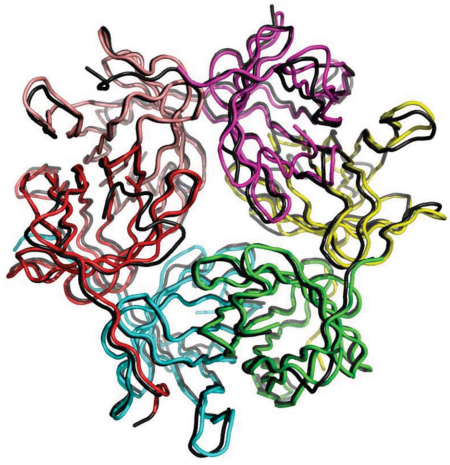
2. Materials and methods

The expression, purification and crystallization of rhizavidin in apo and biotin-complexed forms were performed as described previously (Meir *et al.*, 2009). Apo rhizavidin was crystallized using the sitting-drop method, in which the reservoir solution contained 20–24% PEG 4000, 0.1 *M* Tris–HCl buffer pH 7.8–8.5, 0.2 *M* $MgCl_2$ and 5 m*M* DTT. Prior to data collection, crystals were cryoprotected in 25% ethylene glycol; they belonged to the orthorhombic space group $P2_12_12$ (unit-cell parameters $a = 115.8$, $b = 116.8$, $c = 48.9$ Å). Diffracting crystals of the biotin complex were grown using the sitting-drop vapour-diffusion method using somewhat different conditions compared with the apo form: the reservoir solution contained 20–22% PEG 6000, 0.1 *M* bis-tris buffer pH 6–6.5, 0.15 *M* $CaCl_2$, 10 m*M* sodium pyrophosphate and 3 m*M* DTT. Crystals were cryoprotected using 25% glycerol and belonged to the orthorhombic space group $C222_1$ (unit-cell parameters $a = 45.7$, $b = 130.0$, $c = 237.6$ Å).

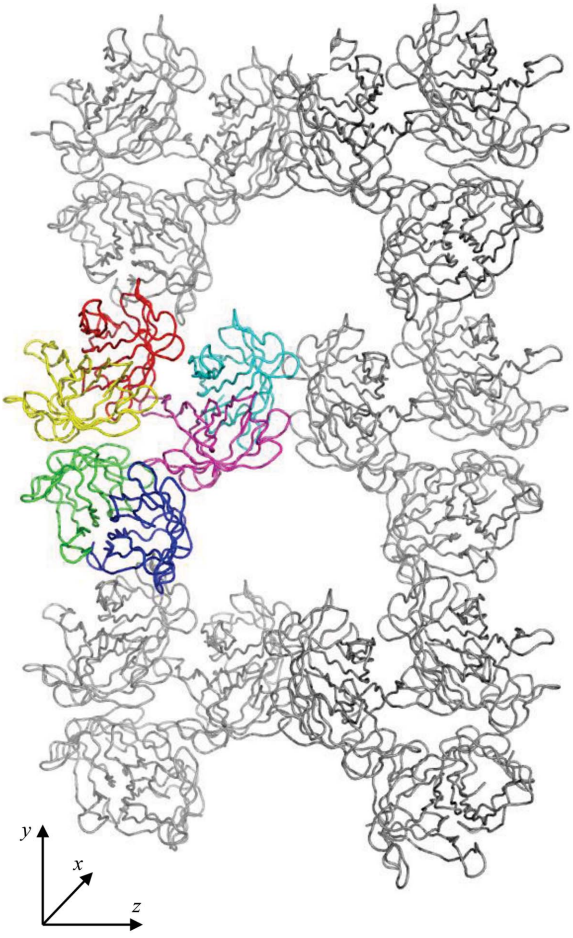
The structure of apo rhizavidin was determined as described previously *via* SAD using an SeMet derivative and was refined to 1.5 Å resolution using native data (Table 1; Meir *et al.*, 2009). The asymmetric unit included six rhizavidin monomers organized as a trimer of dimers to form a bagel-shaped hexameric assembly (Figs. 2*a* and 3*a*). The Matthews coefficient (Matthews *et al.*, 1974) suggested the presence of 4–6 copies of rhizavidin in the asymmetric unit, but the SAD solution clearly indicated six monomers with 40% crystal solvent content and $V_M = 2.03$ Å^3 Da^{-1}.

In the solution of the biotin complex *via* molecular replacement, we reasonably assumed similar packing in the asymmetric unit as in the apo hexameric assembly. However, the molecular-replacement search using *AMoRe* (Navaza, 1994) implemented in *CCP*4*i* (Potterton *et al.*, 2003) with the hexamer as a search model resulted in an unlikely solution that displayed total molecular overlaps when generating the symmetry-related molecules (Fig. 2*a*). The solution appeared to be incorrect and the molecular-replacement search was pursued using a rhizavidin dimer or monomer as a search model. Using *AMoRe* (Navaza, 1994) with a monomer in different resolution ranges did not result in a viable solution. In the case of the dimer, the rotation/translation results also could not provide a viable solution. There were no apparent peaks that could indicate a solution.

When using *MOLREP* (Vagin & Isupov, 2001) with a monomer as a search model in the resolution ranges 15.0–5.0, 15.0–4.5 and 15.0–4.0 Å no solution could be generated (probably owing to the fraction of scattering matter that the model represented). Using a rhizavidin dimer as the search model resulted in a molecular-replacement solution consisting of three pairs, which appeared to be arranged



(*a*)



(*b*)

Figure 2
Crystal-packing diagrams of the molecular-replacement solutions of the rhizavidin–biotin complex. (*a*) The molecular-replacement solution using the apo rhizavidin hexamer as the search model resulted in a solution, shown in different colours for each monomer, in which the symmetry-related molecule (shown as black tubes) overlapped completely with the solution. (*b*) The molecular-replacement solution of three rhizavidin dimers shown in tube representation in individual colours. One layer of symmetry-related molecules form the molecular packing, which displays a significant void which could accommodate one or two rhizavidin monomers. This observation as well as the refinement parameters propagated a search for additional rhizavidin molecules in the asymmetric unit.

differently from those observed for apo rhizavidin (Fig. 2*b*; R = 49.5%, score = 45.0% in the resolution range 20–4.5 Å). Accordingly, there were three strong rotation solutions with a prominent signal-to-noise ratio. When generating the symmetry-related molecules, a void was observed which could accommodate one or two molecules (Fig. 2*b*). This solution was refined by *REFMAC*5 (Murshudov *et al.*, 1999) in the resolution range 50–2.3 Å, resulting in an R value of 35.3% and $R_{\rm free}$ = 40.8%. The electron-density map in the packing void was not clear and did not indicate any secondary-structure features that might hint at the orientation of the additional molecule(s).

In order to resolve the missing structural segment we resorted again to *AMoRe*. We used the six monomers (Fig. 2*b*) found by *MOLREP* as the fixed model (peak 1 in the rotation search) and a rhizavidin dimer as the second search model (peak 8 in the rotation search). The solution filled in the 'void', but there were two overlapping molecules as observed in the initial search. The overlapping molecules were attributed to the additional dimer and we thus elected to remove one of them, which resulted in acceptable crystallographic packing. The resultant R value was 40.8% and the correlation coefficient was 52.8% in the resolution range 15.0–4.5 Å. The solved model contained seven rhizavidin monomers in the asymmetric unit with a relatively low solvent content of 30% and a $V_{\rm M}$ of 1.74 Å^3 Da^{-1} (Fig. 4*a*). Consequently, rigid-body and constrained refinement protocols in *REFMAC*5 were conducted, resulting in an R value of 25.5% and an $R_{\rm free}$ of 31.3% in the resolution range 50–2.3 Å prior to model building (Table 1; Meir *et al.*, 2009).

3. Results and conclusions

Although rhizavidin inherently consists of the 1–4 dimer, the structural configurations of rhizavidin crystals in the apo and biotin-complexed forms are intriguing; they consist of six and seven monomers, respectively. These structural arrangements are stabilized

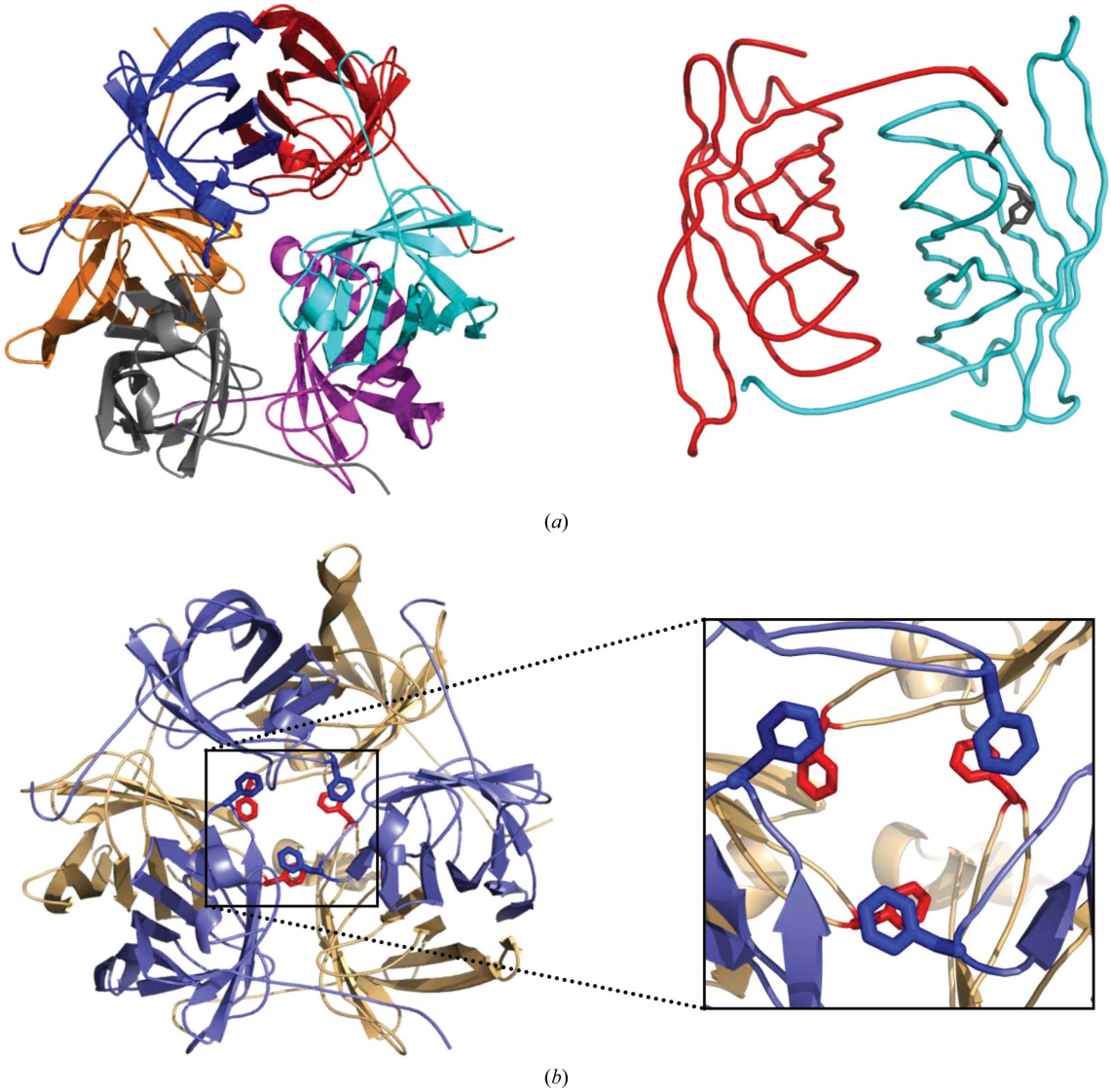


Figure 3
(*a*) Ribbon representation (left) of the asymmetric unit of apo rhizavidin. The hexamer consists of a trimer of dimers and each monomer is shown in a distinctive colour (cyan, red, yellow, magenta, grey and pink). The nonbiological dimer (right) in which the C′-terminus of monomer *A* (red) is perturbed into part of the biotin-binding site of monomer *C* (cyan). Upon biotin binding this region is occupied by the ligand (modelled in black) and cannot accommodate the C-termini. (*b*) The hydrophobic cavity of the rhizavidin hexamer contains six phenylalanine residues (Phe34) contributed by each monomer (blue and red) thus generating two triangles at two parallel planes. An enlargement of the six Phe residues at the core of the hexamer is shown on the right.

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the outer shell.

	Rhizavidin	Rhizavidin–biotin complex
Wavelength (Å)	0.9794	0.9794
ESRF beamline	ID23-1	ID23-1
Space group	$P2_12_12$	$C222_1$
Unit-cell parameters (Å)	a = 115.8, b = 116.8, c = 48.9	a = 45.7, b = 130.0, c = 237.6
Resolution range (Å)	45.0–1.5 (1.55–1.50)	50–2.3 (2.38–2.30)
Unique reflections	104982	31739
Redundancy	4.6	4.9
R_{merge}†	8.1 (66.9)	6.3 (30.0)
Completeness	98.1 (90.6)	98.7 (92.9)
$I/\sigma(I)$	15.2 (1.3)	26.3 (3.2)
No. of protein atoms	6344	6002
No. of ligand atoms	—	112
No. of solvent atoms	717	122
R factor	16.8	21.3
R_{free}‡	22.2	30.4
Average B factor (Å^2)		
Protein	24.6	48.2
Ligand	—	39.9
Solvent	38.1	44.3
R.m.s.d. from ideality		
Bond lengths (Å)	0.014	0.015
Bond angle (°)	1.61	1.74
Ramachandran plot (%)		
Favoured	89.3	85.4
Allowed	10.3	14.1
Generously allowed	0.4	0.5
Disallowed	0.0	0.0

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ The test set consisted of 5% of all data.

by remarkable molecular and crystal contacts which contribute to these unique structural configurations.

As indicated previously, the asymmetric unit in apo rhizavidin forms a bagel-shaped hexameric assembly which is constructed as a trimer of dimers (Fig. 3*a*). The hexameric structure in the crystal is stabilized by several intermolecular interactions. A remarkable contribution to the hexamer assembly consists of an interaction in which the C-terminal segment of one monomer is contributed to a neighbouring monomer which belongs to a nonbiological dimer and *vice versa* in a 'yin and yang'-like configuration (Fig. 3*a*). Intriguingly, the end of the C′-terminus (residues 132–135) protrudes into part of the biotin-binding site of its adjacent monomer and seals its top. The contact surface of the 'yin and yang' dimer is larger than that of the inherent 1–4 dimer, with average values of 1575 and 1314 Å^2, respectively. The inner core formed by the bagel-shaped hexamer is mainly of polar nature with approximately 70 solvent molecules accommodating its space. Both ends of the inner core are sealed by part of the N-termini segments of each of the monomers. One of the most notable elements is the presence of six phenylalanine residues (Phe34), which forms interactions between non-inherent monomers in the hexameric core. In this arrangement three Phe34 residues are located at opposite sides of the core to form two almost perfect equilateral and parallel triangles 18.8 Å apart (Fig. 3*b*). The Phe34 residues in each triangle are too distant (∼9 Å) to interact with each other, but form interactions with alternating monomers in the hexameric assembly.

The asymmetric unit of the rhizavidin–biotin complex crystal consists of seven monomers. This fascinating heptamer (Fig. 4) is constructed of three and a half rhizavidin pairs. The heptamer in the asymmetric unit consists of two independent rhizavidin dimers and three monomers, forming one and a half dimers which could also be regarded as half a hexamer. This crystal packing substantiates the presence of rhizavidin dimers, as also observed in solution. The unpaired trimer in the asymmetric unit attains its counterpart from a symmetry-related molecule by a twofold screw operation along the y axis to form a bagel-shaped hexamer similar to that observed in the apo structure (Fig. 4). At the core of the rhizavidin hexamer there are similar molecular elements (*i.e.* Phe34) as observed for the apo packing. Since the N′-termini are less ordered than in the apo form (there is undefined electron density in this region) it is probably accommodated differently than in the apo model. Consequently, the ends of the inner molecular core of the hexamer are more available to solvent. In addition, the C′-termini of rhizavidin do not occupy the edge of the biotin-binding site and are shifted to a position in its vicinity.

This highly compact and aesthetic hexameric arrangement promoted re-examination of the oligomeric state of rhizavidin in solution. Size-exclusion chromatography of rhizavidin verified its inherent dimeric configuration under different conditions (protein concentration, pH, ionic strength) in both its apo and biotin-complexed forms without any indication of higher oligomeric structures (data not shown).

Despite the high overall similarity in tertiary structure to avidin and streptavidin, rhizavidin exhibits some intriguing differences, mainly in its dimeric arrangement, while maintaining its high affinity towards biotin. Another unique feature which appears to be more than a fluke is its capability to form unique bagel-like hexamers in the crystalline state. These novel oligomerization states that were not

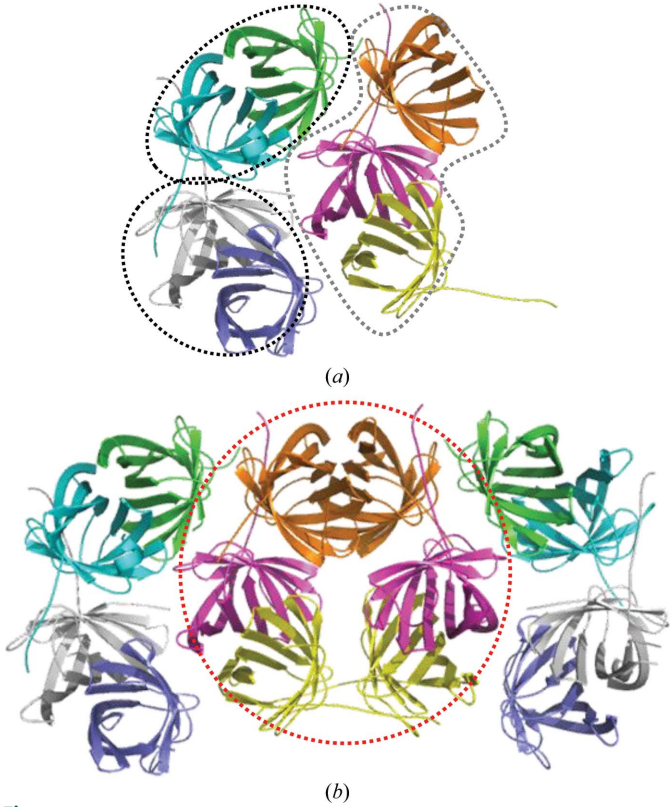


Figure 4
The molecular-replacement solution (*a*) of the rhizavidin–biotin complex displays seven monomers in the asymmetric unit, which are shown in individual colours. The two dotted ellipses indicate the individual 1–4 rhizavidin dimers. The other dotted structure indicates the three monomers; the inherent dimer is shown in magenta and yellow and the unmatched monomer in orange. When generating the symmetry-related operation (*b*) it is apparent that the latter trimeric arrangement is half a hexamer (indicated by the red dotted circle), similar to that observed for the apo structure.

observed previously established both the biochemical properties and the nontrivial structure solution of rhizavidin. The novel oligomerization forms of avidin-like proteins may lead to novel and innovative research in protein design and biotechnology.

References

Bayer, E. A. & Wilchek, M. (1990). *J. Chromatogr.* **510**, 3–11.

Dehlinger, D. A., Sullivan, B. D., Esener, S. & Heller, M. J. (2007). *Small*, **3**, 1237–1244.

DeLano, W. L. (2002). *The PyMOL Molecular Viewer*. http://www.pymol.org.

Freitag, S., Le Trong, I., Chilkoti, A., Klumb, L. A., Stayton, P. S. & Stenkamp, R. E. (1998). *J. Mol. Biol.* **279**, 211–221.

Green, N. M. (1975). *Adv. Protein Chem.* **29**, 85–133.

Green, N. M. (1990). *Methods Enzymol.* **184**, 51–67.

Helppolainen, S. H., Nurminen, K. P., Määttä, J. A., Halling, K. K., Slotte, J. P., Huhtala, T., Liimatainen, T., Ylä-Herttuala, S., Airenne, K. J., Närvänen, A., Jänis, J., Vainiotalo, P., Valjakka, J., Kulomaa, M. S. & Nordlund, H. R. (2007). *Biochem. J.* **405**, 397–405.

Hendrickson, W. A., Pahler, A., Smith, J. L., Satow, Y., Merritt, E. A. & Phizackerley, R. P. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 2190–2194.

Kurzban, G. P., Bayer, E. A., Wilchek, M. & Horowitz, P. M. (1991). *J. Biol. Chem.* **266**, 14470–14477.

Laitinen, O. H., Airenne, K. J., Marttila, A. T., Kulik, T., Porkka, E., Bayer, E. A., Wilchek, M. & Kulomaa, M. S. (1999). *FEBS Lett.* **461**, 52–58.

Laitinen, O. H., Hytönen, V. P., Nordlund, H. R. & Kulomaa, M. S. (2006). *Cell. Mol. Life Sci.* **63**, 2992–3017.

Laitinen, O. H., Nordlund, H. R., Hytönen, V. P. & Kulomaa, M. S. (2007). *Trends Biotechnol.* **25**, 269–277.

Livnah, O., Bayer, E. A., Wilchek, M. & Sussman, J. L. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 5076–5080.

Määttä, J. A., Airenne, T. T., Nordlund, H. R., Jänis, J., Paldanius, T. A., Vainiotalo, P., Johnson, M. S., Kulomaa, M. S. & Hytönen, V. P. (2008). *Chembiochem*, **9**, 1124–1135.

Matthews, B. W., Weaver, L. H. & Kester, W. R. (1974). *J. Biol. Chem.* **249**, 8030–8044.

Meir, A., Helppolainen, S. H., Podoly, E., Nordlund, H. R., Hytönen, V. P., Määttä, J. A., Wilchek, M., Bayer, E. A., Kulomaa, M. S. & Livnah, O. (2009). *J. Mol. Biol.* **386**, 379–390.

Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Paganelli, G., Bartolomei, M., Ferrari, M., Cremonesi, M., Broggi, G., Maira, G., Sturiale, C., Grana, C., Prisco, G., Gatti, M., Caliceti, P. & Chinol, M. (2001). *Cancer Biother. Radiopharm.* **16**, 227–235.

Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. (2003). *Acta Cryst.* D**59**, 1131–1137.

Sano, T., Vajda, S., Smith, C. L. & Cantor, C. R. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 6153–6158.

Urbano, N., Papi, S., Ginanneschi, M., De Santis, R., Pace, S., Lindstedt, R., Ferrari, L., Choi, S., Paganelli, G. & Chinol, M. (2007). *Eur. J. Nucl. Med. Mol. Imaging*, **34**, 68–77.

Vagin, A. A. & Isupov, M. N. (2001). *Acta Cryst.* D**57**, 1451–1456.

Weber, P. C., Cox, M. J., Salemme, F. R. & Ohlendorf, D. H. (1987). *J. Biol. Chem.* **262**, 12728–12729.

Wilchek, M. & Bayer, E. A. (1990). *Methods Enzymol.* **184**, 5–13.