

# Identification of novel target sites and an inhibitor of the dengue virus E protein

Ragothaman Yennamalli · Naidu Subbarao ·  
Thorsten Kampmann · Ross P. McGeary ·  
Paul R. Young · Bostjan Kobe

Received: 17 November 2008 / Accepted: 7 February 2009 / Published online: 25 February 2009  
© Springer Science+Business Media B.V. 2009

**Abstract** Dengue and related flaviviruses represent a significant global health threat. The envelope glycoprotein E mediates virus attachment to a host cell and the subsequent fusion of viral and host cell membranes. The fusion process is driven by conformational changes in the E protein and is an essential step in the virus life cycle. In this study, we analyzed the pre-fusion and post-fusion structures of the dengue virus E protein to identify potential novel sites that could bind small molecules, which could interfere with the conformational transitions that mediate the fusion process. We used an *in silico* virtual screening approach combining three different docking algorithms (DOCK, GOLD and FlexX) to identify compounds that are likely to bind to these sites. Seven structurally diverse molecules were selected to test experimentally for inhibition of dengue virus propagation. The best compound showed an  $IC_{50}$  in the micromolar range against dengue virus type 2.

**Keywords** Dengue virus · Envelope protein · Flavivirus · Virtual drug screening

## Abbreviations

BSA	Bovine serum albumin
DENV	Dengue virus
DMEM	Dulbecco's modified Eagle's medium
FCS	Fetal calf serum
JEV	Japanese encephalitis virus
MTT	Methyl thiazole tetrazolium
PDB	Protein Data Bank
PBS	Phosphate buffered saline
pfu	Plaque-forming units
TBEV	Tick-borne encephalitis virus
WNV	West Nile virus

## Introduction

Dengue viruses (DENV) are members of the flaviviruses, a group of some 70 viruses, many of which (including West Nile virus (WNV), yellow fever virus, Japanese encephalitis virus (JEV), tick-borne encephalitis (TBEV) and Australian encephalitis viruses) are associated with human disease. DENV are a major public health problem in tropical and sub-tropical areas with up to 100 million people infected annually [1]. Infection with this virus can result in acute febrile illness (dengue fever) and in severe cases leads to abnormalities in vascular permeability and haemostasis (dengue haemorrhagic fever and dengue shock syndrome). DENV are small, enveloped viruses with a positive-sense RNA genome that encodes ten viral proteins. Mature dengue virions contain three of these proteins: a capsid protein (C), associated with genomic

**Electronic supplementary material** The online version of this article (doi:10.1007/s10822-009-9263-6) contains supplementary material, which is available to authorized users.

R. Yennamalli · N. Subbarao  
Centre for Computational Biology and Bioinformatics, School  
of Information Technology, Jawaharlal Nehru University,  
New Delhi 110067, India

R. Yennamalli · T. Kampmann · R. P. McGeary ·  
P. R. Young · B. Kobe (✉)  
School of Chemistry and Molecular Biosciences, University  
of Queensland, Brisbane, QLD 4072, Australia  
e-mail: b.kobe@uq.edu.au

P. R. Young · B. Kobe  
Institute for Molecular Bioscience, University of Queensland,  
Brisbane, QLD 4072, Australia

RNA, and two integral membrane proteins, the membrane protein (M) and the envelope protein (E).

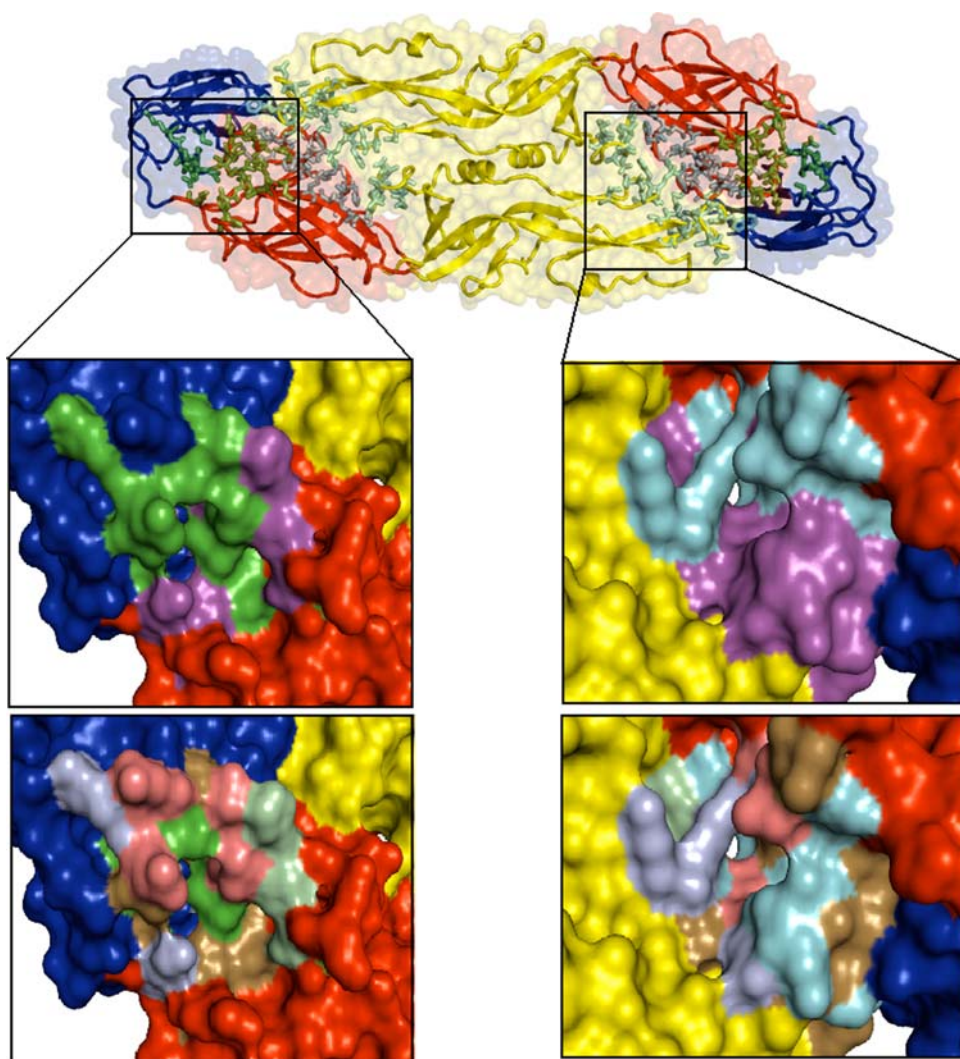
The entry of enveloped viruses into host cells requires the attachment to host cell receptors, followed by fusion of the viral envelope with cellular membranes. For flaviviruses, the glycosylated E protein (Fig. 1) mediates both receptor binding and fusion activities [2, 3]. Infectious entry occurs within the endosome, following uptake of intact virus particles through receptor-mediated endocytosis. Protonation of histidines in the low-pH environment of the endosome is the likely trigger for the conformational change that leads to the reversible dissociation of virion surface E dimers into monomers, followed by an irreversible transition to homotrimers [4–6]. A hydrophobic stretch of amino acids, referred to as the “fusion peptide”, is exposed during this conformational change and interacts with the target membrane, initiating the fusion process and ultimately the merging of the two bilayers [7].

The DENV E protein comprises a large N-terminal ectodomain, which is connected via a short “stem” region to

two membrane-spanning domains that anchor the protein in the lipid bilayer. The crystal structures of pre- and post-fusion forms of a truncated version of the E ectodomain (DENV E protein minus the stem and anchor regions) have been determined for both DENV (Fig. 1) and TBEV [5, 8–10]. Each monomer of the crystallized dimeric E protein contains three domains (I–III; Fig. 1). One of the initial structural changes in the E protein induced by exposure to low pH is a movement around a hinge-like region separating domains I and II (Fig. 1), which is thought to expose the internal fusion peptide for interaction with the opposing cell membrane. In addition, there is a radical shift of domain III into the interface between domains I and II [5]. These structural transitions are obvious targets for inhibitor design.

The current method to control the spread of disease caused by DENV is to control the vector, the mosquito *Aedes aegypti*. The only available disease treatment is supportive therapy. For these reasons, antiviral agents have been sought intensively; a number of approaches have been reported, including nucleoside triphosphate synthesis

**Fig. 1** Ribbon diagram of the pre-fusion dimeric structure of the ectodomain of DENV2 E protein (PDB ID: 1OKE) [8]. The residues of the two sites selected as suitable for targeting with compounds are shown in stick representation, site 1 in green and site 2 in cyan. The surface of the E protein is also shown in semi-transparent representation. Domain I is coloured red, domain II is coloured yellow and domain III is coloured blue. The insets highlight sites 1 and 2. *Top left*, site 1 shown in surface representation; residues conserved in at least 80% of DENV sequences are highlighted in magenta. *Bottom left*, site 1 shown in surface representation, highlighting charged residues (Asp and Glu in pink, Lys and Arg in light blue), histidines (in lime green) and hydrophobic residues (Val, Leu, Ile, Phe, Tyr, Pro, Met, Cys, Thr in sand). *Top right*, site 2 shown in surface representation; residues conserved in at least 80% of DENV sequences are highlighted in magenta. *Bottom right*, site 2 shown in surface representation, highlighting amino acid types as in *bottom left*



inhibitors such as mycophenolic acid, ribavirin and 6-azauridine [11], suppressors of viral RNA synthesis such as triaryl pyrazoline [12], nucleoside analogues [13] and tetra-aldehyde [14] inhibitors of NS3 protein helicase and protease activities, peptides that mimic the conserved protein cleavage sites [15], host alpha-glucosidase inhibitors suppressing virion secretion and infectivity such as castanospermine [16, 17] and deoxynojirimycin [18], c-Src protein kinase inhibitors that inhibit the assembly and maturation of DENV [19], monoclonal antibodies [20–23], domain III of the fusion protein [24], RNA silencing [25–27], peptides targeting the E protein [28, 29], and polyanions preventing host cell receptor binding [30].

Virtual (in silico) screening of chemical libraries is emerging as a powerful approach to identify lead compounds. Although this approach requires knowledge of the three-dimensional structure of the target protein, and various necessary assumptions and approximations currently limit the success rates of most implementations, it has been used successfully in a number of systems [31–33]. The crystal structure of the ectodomain of DENV type 2 (DENV2) revealed a hydrophobic pocket in the hinge region between domains I and II, which was proposed as a suitable target site for small-molecule inhibitors of the fusion process [8]. Recently, this pocket has been targeted using virtual screening, and tetracycline derivatives with antiviral inhibitory activity have been identified [34]. We examined the crystal structures of DENV E protein in pre- and post-fusion states to identify alternative target sites for small molecule binding and inhibition of the fusion process. We then used virtual screening combining three different docking algorithms (DOCK, GOLD and FlexX) with a number of chemical databases to identify suitable compounds predicted to bind in these sites. Seven structurally diverse potential inhibitors were then selected for testing in a cell culture-based antiviral assay, and one was found to significantly reduce viral replication with an  $IC_{50}$  values of  $\sim 4 \mu M$ .

## Materials and methods

### Identification of novel binding sites

Cavity-detection algorithm “Putative active site with spheres” (PASS) [35] was used to identify cavities in the proteins. Consurf [36] was used to find the residues that are conserved among the four DENV serotypes. Cavities were selected if either (1) the cavity was present in the dimer but not in the trimer, or (2) the cavity was present in the dimer interface. The spatial arrangement of E protein transmembrane region and the presence of M protein in the mature virion make only one side of the E protein accessible for a compound to bind; only cavities on the receptor-facing side

of the E protein were therefore considered [37]. Suitable cavities were checked further based on the following criteria [38]: (1) functionality—an inferred role in the conformational transitions; (2) presence of hydrophobic residues; (3) presence of charged residues; (4) solvent accessibility; and (5) absence of carbohydrate moieties (within a radius of  $10 \text{ \AA}$ ).

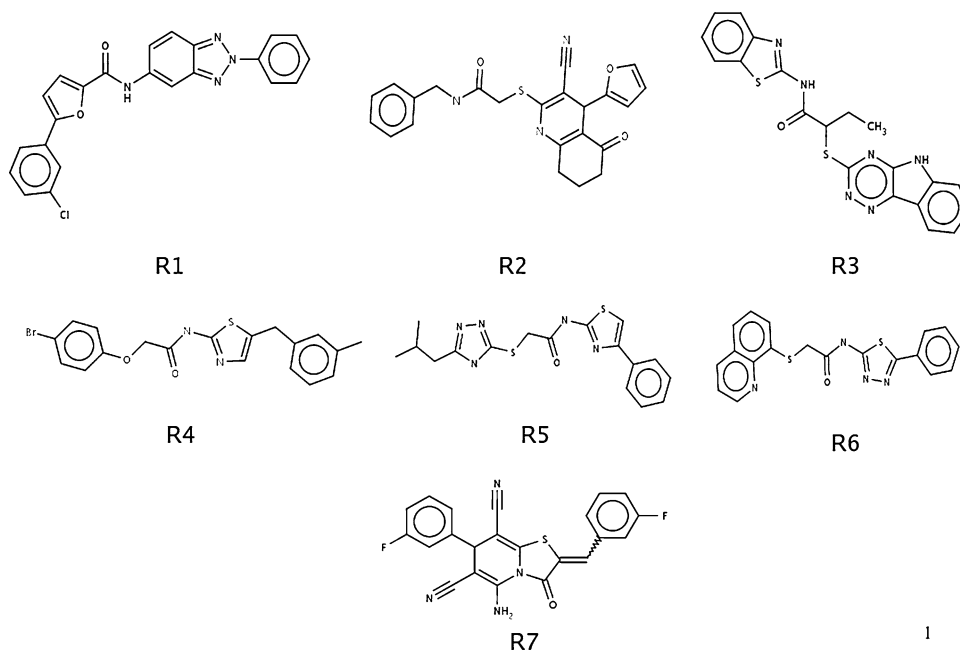
### Virtual screening with the dengue virus E protein

The workflow of our virtual screening procedure is shown schematically in Supplementary Fig. 1. Thirteen publicly and commercially available databases were accessed (Supplementary Table 1). The databases were filtered with the drug-likeness-index [39] using Filter v1.0.2 (OpenEye Scientific Software). Virtual screening was performed on this subset of molecules with three docking algorithms, GOLD v2.1.2 [40], FlexX v2.2.0 [41] and DOCK v5.1.1 [42]. The molecules were converted to three-dimensional structure files using CORINA v2.64 (Molecular Networks GmbH Computerchemie). Hydrogen atoms were added explicitly using Babel v1.6 (CCL.NET). The parameters used for DOCK are listed in Supplementary Tables 2 and 3. In FlexX, the active site was defined in the receptor descriptor file and the surface was generated. The ligand was fragmented into flexible and rigid parts using the “automatic” mode. The rigid parts were placed in the active site using the triangle algorithm method. The docking of the ligand was then completed using the command “complex all”. In GOLD, the molecules were first screened using the “library screening” setting; 10% of the top-ranking molecules were taken for a second round of screening using the “3-times-up” setting; and finally, the top-scoring molecules were screened using the standard default setting. The parameters used for the different settings are listed in Supplementary Table 4. As GOLD uses a genetic algorithm approach, which is a heuristic method, ten independent docking runs were performed for each molecule.

A molecule was ranked higher if it scored well with all three different methods. The methods have different search algorithms and scoring functions. Hence, it was not possible to compare the fitness scores of GOLD, FlexX and DOCK directly. We instead used the X-Score v1.2.1 scoring function [43]. X-Score calculates the negative logarithm of dissociation constant of the ligand to the protein,  $-\log(K_d)$ , as the average of three scoring functions (HPScore, HMScore and HSScore), and predicts the binding energy (kcal/mol) of the ligand. X-Score was reported to have an accuracy of  $\pm 2.2$  kcal/mol to the actual binding energies [43]. Hence, the molecules that had the binding energies of the docked conformations obtained by the three methods within the  $\pm 2.2$  kcal/mol range of each other were considered for further analysis.

We used GOLD as a reference to compare the results of DOCK and Flex-X [44]. Molecules with a GOLD fitness score 70 and above [40] were selected to analyze interactions in the binding site using Ligplot v4.4 [45]. These molecules were clustered according to sub-structure similarity using LibMCS of Jchem v3.2.11 (ChemAxon) to yield a diverse set of molecules. Molecules were further assessed for stability, ease of synthesis and modification of chemical groups, and structural diversity. Seven molecules (R1–R7; Fig. 2) were finally selected for further computational and experimental testing. Assessments of stabilities and ease of syntheses were based on the judged likelihood of the candidate compounds undergoing detrimental oxidation or elimination reactions, the number of chiral centres in the molecules, the functional groups within the molecules (preferably amide linkages), and the availability of the constituent building blocks. As the compounds are commercially available, they are obviously reasonably stable with regards to synthesis, purification, storage and shipping. A number of the rejected compounds contained a cyclohexene-diol, which might be expected to lead to dehydration and aromatisation under the conditions of the assay. Other compounds contained a thiol, which could lead to oxidation to form a disulfide, or an dialkyl sulfide, which could lead to oxidation to a sulfoxide. Many of the rejected candidates had multiple chiral centres, which would make synthesis and purification unattractive. The final seven compounds have either none, or just one chiral centre. Several compounds were rejected as they did not contain amide bonds. All of the final seven compounds (except R7) contain amide bonds, which are easy to synthesize.

**Fig. 2** The seven compounds identified by virtual screening against site 1 that were selected for experimental testing



#### Virtual screening with the envelope proteins of other flaviviruses

The seven molecules (R1–R7; Fig. 2) were also docked using GOLD v2.1.2 (standard default settings) and FlexX v2.2.0 to the crystal structures of the envelope proteins from TBEV (Protein Data Bank [PDB] ID: 1SVB) [9], WNV (PDB ID: 2HG0) [46], Japanese encephalitis virus [47] and DENV type 3 (DENV3; PDB ID: 1UZG) [48]. The interactions were compared as described above using X-Score v1.2.1 and analyzed using Ligplot v4.4.

#### Experimental testing of compounds for antiviral activity

##### Cell and virus culture

Vero (African green monkey kidney epithelial) cells (ATCC) were grown in Dulbecco's modified Eagle's medium (DMEM; Invitrogen) supplemented with 3% fetal calf serum (FCS) at 37 °C and 5% CO<sub>2</sub>. FCS was heat-inactivated at 56 °C for 20 min. The cells were subcultured frequently when they reached 90% confluence. DENV2 strain New Guinea C (NGC) stock was stored as 1 mL aliquots at –80 °C.

##### Preparation of compounds

Seven compounds (R1–R6 obtained from Chemdiv Inc. ([www.chemdiv.com](http://www.chemdiv.com)), R7 obtained from Ambinter SARL ([www.ambinter.com](http://www.ambinter.com))) were solubilized in 100% dimethyl sulfoxide (DMSO) to a final concentration of 8 mM. These



were filter-sterilized using a 22  $\mu\text{m}$  filter and stored at  $-20\text{ }^{\circ}\text{C}$ .

#### *Cytotoxicity assay*

The methyl thiazole tetrazolium (MTT) viability assay was used to test for cytotoxicity. The compounds were serially diluted in a 96-well plate to concentrations ranging from 200 to 1.5625  $\mu\text{M}$ . Concentrated hydrochloric acid (HCl) and phosphate buffered saline (PBS) were used as positive and negative controls, respectively. Forty-five  $\mu\text{L}$  of dilutions were added to a Vero cell monolayer and incubated for 2 h at  $37\text{ }^{\circ}\text{C}$  and 5%  $\text{CO}_2$ . Ten  $\mu\text{L}$  of freshly prepared filtered-sterilized solution of MTT (5 mg/mL) was then added to the cells and they were further incubated at  $37\text{ }^{\circ}\text{C}$  and 5%  $\text{CO}_2$  for 4 h, after which 150  $\mu\text{L}$  of lysing solution (0.04 M HCl in 2-propanol) was added. The plates were analyzed in a plate reader (Labsystems MultiskanEX) at 570 and 630 nm, and the difference in values between 570 and 630 nm was plotted as a function of the concentration of the compound.

#### *Plaque reduction assay*

A plaque assay was used to measure antiviral activity. A day before the assay, Vero cells were subcultured and incubated at  $37\text{ }^{\circ}\text{C}$  overnight in a 96-well micro-titre plate to form a monolayer. The compounds were serially diluted in several steps using 100% DMSO from their initial concentration of 8 to 1.5625  $\mu\text{M}$ , maintaining the same DMSO concentrations in the serial dilutions. At the time of introducing the mixture of virus and diluted compounds, the DMSO concentrations were diluted from 100 to 2%. The virus was added to the compounds and incubated for half an hour at room temperature. Forty-five  $\mu\text{L}$  of the mixture of virus and diluted compounds was used to infect the cells. After 2 h at  $37\text{ }^{\circ}\text{C}$  and 5%  $\text{CO}_2$ , the monolayer was overlaid with 150  $\mu\text{L}$  DMEM media containing 3% carboxymethyl cellulose (CMC), Medium 199 (Invitrogen) and 5% FCS, and the plates were incubated for 5 days. The media overlay of the plates was removed by washing three times with PBS. The cells were then fixed using 200  $\mu\text{L}$  of ice-cold 80% acetone in PBS for 20 min at  $-20\text{ }^{\circ}\text{C}$ . The cells were left to dry overnight after removing the fixative. When dried, 100  $\mu\text{L}$  of a mixture of 90% methanol and 10% hydrogen peroxide was added and left for 20 min in the dark at room temperature to inactivate any cellular peroxidase activity. The cells were washed twice with 200  $\mu\text{L}$  of PBS containing 0.1% bovine serum albumin (BSA) for 5 min, and blocked for 20 min at room temperature using PBS/0.1% BSA. Fifty  $\mu\text{L}$  of primary antibody (1:500 dilution of rabbit polyclonal anti-NS1 in

PBS/0.1% BSA) was added to the cells and incubated for 1 h at room temperature. The antibodies were removed by washing the cells three times with 200  $\mu\text{L}$  of PBS/0.1% BSA for 5 min. Fifty  $\mu\text{L}$  of secondary antibody mixture (1:200 dilution of peroxidase labelled goat anti-rabbit IgG in PBS/0.1% BSA) was added to the cells and incubated for 1 h at room temperature. The antibodies were removed by washing the cells twice with 200  $\mu\text{L}$  of PBS/0.1% BSA for 5 min. To visualize the plaques, the cells were treated with 50  $\mu\text{L}$  of freshly prepared 3,3'-diaminobenzidine tetrahydrochloride (DAB) solution, following the manufacturer's instructions (Sigma–Aldrich), and incubated in the dark for 40 min. The colour development was arrested by rinsing the plate in water, followed by air-drying. The plaques were counted manually. The concentration of the compound that was able to reduce the number of plaque-forming units by 50% was defined as the  $\text{IC}_{50}$ .

## **Results and discussion**

The essential role of the DENV E protein in the fusion process, its location on the surface of the mature virus and the availability of its crystal structure make it a suitable target for structure-based drug design. Furthermore, the structures of the envelope proteins from different flaviviruses are well conserved, suggesting the possibility that compounds active against several flaviviruses may exist.

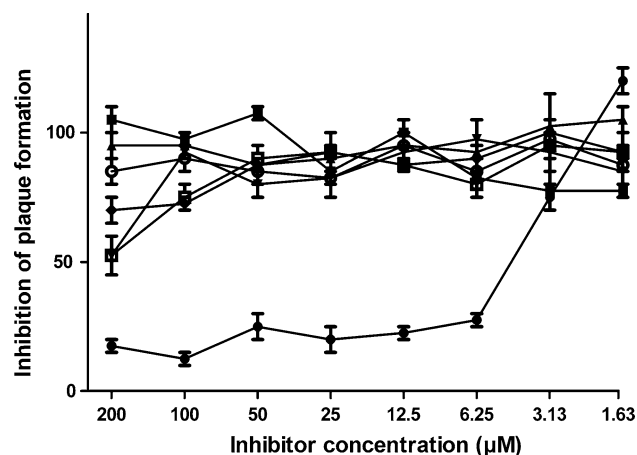
The crystal structure of the ectodomain of DENV2 revealed a hydrophobic pocket in the hinge region between domains I and II, which was proposed to be a suitable target site for small-molecule inhibitors of the fusion process [8]. Recently, this pocket has been targeted using virtual screening and tetracycline derivatives with antiviral inhibitory activity have been identified [34]. We searched for alternative areas in the structure that could be targeted in a similar manner. Suitable cavities were selected if either (1) the cavity was present in the dimer but not the trimer form, and therefore compounds binding to it may inhibit trimer formation; or (2) the cavity was present in the dimer interface, and compounds binding to it may stabilize the dimer or prevent its transition to the post-fusion trimeric form. Six potential sites on the receptor-facing side of the E protein were examined further to determine if they were likely to play a role in the conformational transitions, if they contained both hydrophobic and charged residues, if they were-solvent accessible and if there were no carbohydrate moieties nearby (within a radius of 10 Å) that could interfere with the binding of candidate ligands.

Two sites were finally selected (Fig. 1). Site 1 is present in the pre-fusion structure and there is no corresponding site in the post-fusion structure; it is located between domains I and III. These domains move relative to each

other between the pre- and post-fusion structures. The cavity is made up from 25 residues: residues 1, 143–149, 156, 158, 178 and 295 from domain I, and residues 324, 333, 355–357, and 359–366 from domain III. Among these are six and ten charged and hydrophobic residues, respectively. The cavity has a volume of  $269 \text{ \AA}^3$  and solvent accessible area of about  $936 \text{ \AA}^2$ . Eleven residues that line this pocket are conserved in 80% or more of DENV sequences. The site is also structurally conserved in other flavivirus E proteins. The nearest glycosylated residue (Asn153) is over  $12 \text{ \AA}$  away.

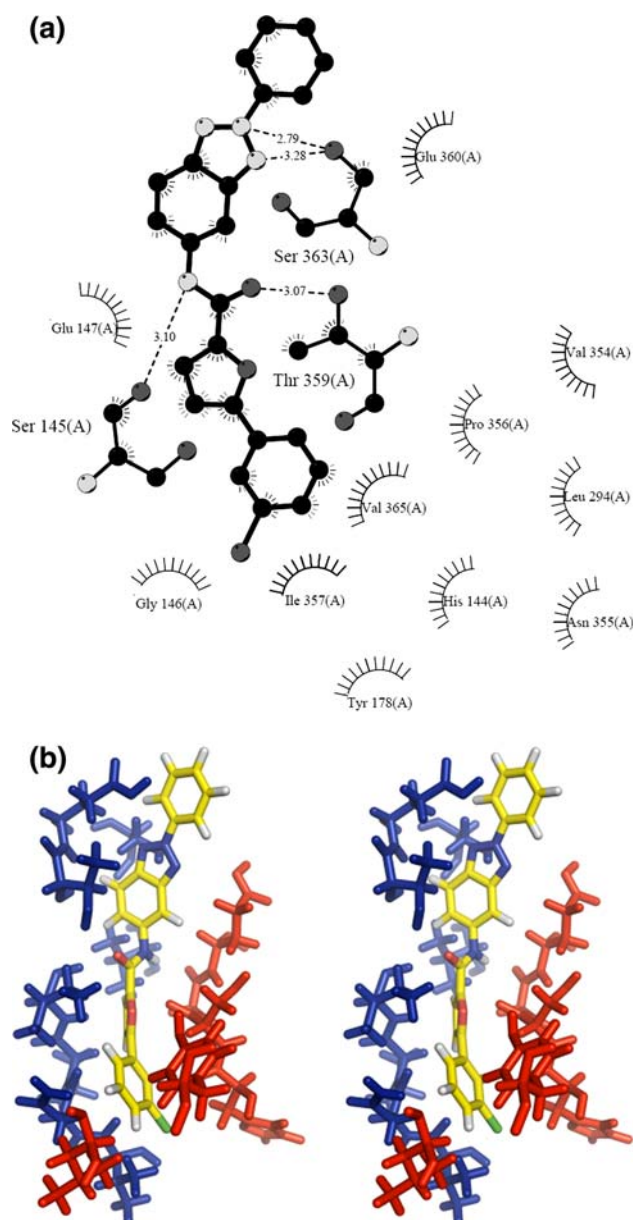
Site 2 lies near the fusion peptide and is also present only in the dimer form. It lies between two subunits of the E protein dimer. The cavity is made up of 33 residues: residues 1–8, 28, 30, 44, 151–155 and 316 of chain A, and 97–109 and 244–247 of chain B. Among these are 7 and 12 charged and hydrophobic residues, respectively. The cavity has a volume of about  $298 \text{ \AA}^3$  and solvent accessible area of  $1,190 \text{ \AA}^2$ . Twenty residues out of the total 33 that line this site are conserved in 80% or more of DENV sequences. The cavity is also conserved in the pre-fusion structures of E proteins from TBEV and JEV. The nearest glycosylated residues (Asn67 and Asn153) are over  $15 \text{ \AA}$  away.

While both sites appear suitable for inhibitor design, Site 1 was targeted first by virtual screening. Thirteen databases of compounds were selected and filtered for compounds that satisfy Lipinski's rule-of-five and are likely to be non-toxic to mammalian cells (Supplementary Fig. 1; 641 627 molecules were filtered to select 114 109 molecules). Eighty molecules had GOLD fitness scores above 70, and 41 were selected based on X-Score analysis



**Fig. 3** Antiviral activity of compounds R1–R7. The graph shows plaque-forming units (pfu) of DENV2 as a percentage of mock-treated controls, as a function of the compound concentration. Each point represents the mean of the values from two independent experiments with *error bars* showing standard deviations. The assay was performed in duplicate. *Symbols* refer to: R1, *closed circles*; R2, *closed squares*; R3, *closed triangles*; R4, *inverted closed triangles*; R5, *closed diamonds*; R6, *open circles*; R7, *open squares*

of DOCK and Flex-X results (see “[Materials and Methods](#)”; Supplementary Table 5). The molecules were further analyzed by substructure clustering with LibMCS, and compounds were removed if thought likely to be unstable or if they were expected to require more than five steps to be synthesized. The H–N–C=O motif was observed to be common among the selected molecules.



**Fig. 4** The predicted interactions of compound R1 with E protein based on computational docking. **a** Schematic diagram of the interactions. Polar contacts are shown with *dashed lines*, and hydrophobic contacts are indicated by *arcs with radiating spokes*. Carbon, nitrogen and oxygen atoms are shown in *black, white and grey*, respectively. Prepared with the programme LIGPLOT [45]. **b** Stereo-diagram of the structure of R1 (*yellow*) and the DENV2-interacting residues (*blue*, domain III and *pink*, domain I), shown in stick representation

The predicted interactions with the protein were inspected further for 31 compounds. Sixteen molecules made at least one hydrogen bond with any one of the ten conserved residues in the pocket. Fourteen molecules had at least one hydrophobic interaction with at least two of the ten conserved residues. The conserved residues Ser145, Tyr178, Pro356, Gly146 and His144 contacted the ligands most often.

Considering commercial availability, seven molecules were finally selected for experimental testing using the plaque assay (Fig. 2; Supplementary Table 6). The compound R1 showed a reduction in the number of plaque-forming units (pfu) and is estimated to have an  $IC_{50}$  of approximately 4  $\mu$ M (Fig. 3). Several controls were used in the assay: (1) virus without compounds added (positive control for virus); (2) media without compounds added (negative control for virus); and (3) DMSO and virus

(negative control for compounds). None of the compounds showed cytotoxicity at 200  $\mu$ M according to the MTT assay, indicating that R1 had a selectivity index ( $SI = CC_{50}/IC_{50}$ ) greater than 50.

The compound R1 is predicted to form hydrogen-bonds with residues Ser363, Thr359 and Ser145 and further interactions with 14 other residues (Fig. 4; Supplementary Table 7). R1 has a high GOLD fitness score and X-Score binding energy (71.72 and  $-9.43$  kcal/mol, respectively) and was the only compound of the seven tested to effect a significant plaque reduction.

The targeted cavity is structurally conserved among other flaviviruses (Fig. 5). Hence, there is a possibility that R1 could be effective against other flaviviral E proteins. R1 was therefore docked against other known flaviviral E protein structures (Fig. 5; Supplementary Table 8). Surprisingly, R1 does not show high GOLD fitness scores and

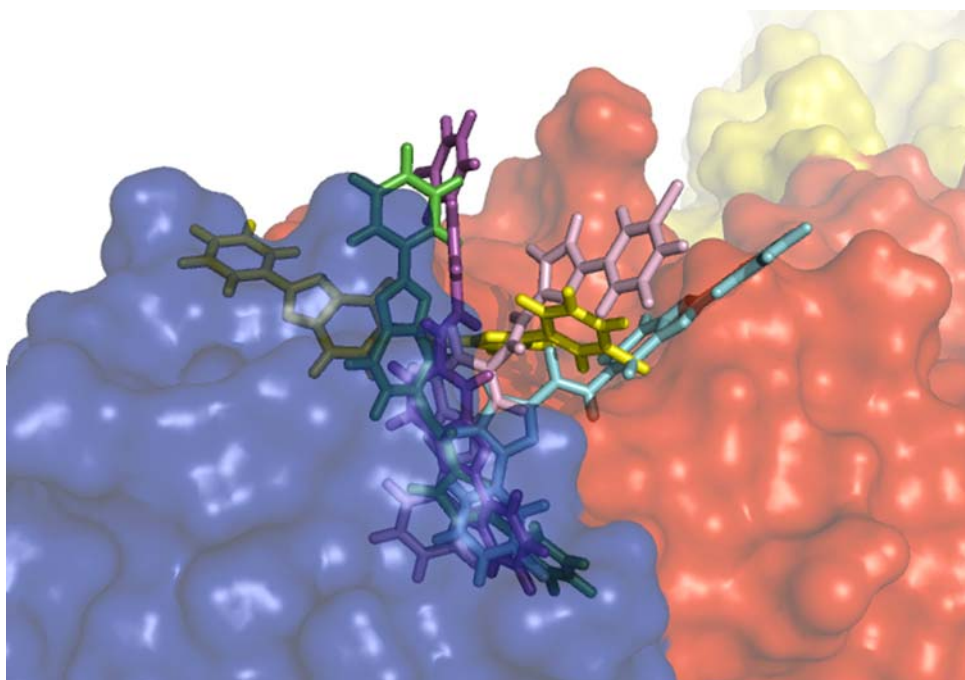
DENV2	MRCIGISNRDFVEGVSGGSWVDIVLEHSGCVTTMAKNKPTLDFELIKTEAKQPATLRKYCIEAKLTNTTT	70
DENV3	MRCVGVGNRDFVEGLSGATWVDVLEHGGCVTTMAKNKPTLDIELQKTEATQLATLRKLCIEGKITNITT	70
WNV	FNCLGMSNRDFLEGVSGATWVDLVLEGDSCTVIMSKDKPTIDVKMMNMEAANLAEVRSYCYLATVSDLST	70
JEV	FNCLGMSNRDFIEGASGATWVDLVLEGDSCLTIMANDKPTLDVRMINIEAVQLAEVRSYCYHASVTDIST	70
TBEV	SRCTHLENRDFVTGTQGTTRVTLVLELGGCVTTITAEKGPSMDVWLDAIYQENPAKTREYCLHAKLSDTKV	70
.....		
DENV2	ESRCPTQGEPTLNEEQDKRFVCKHSMVDRGWNGCGFLFGKGGIVTCAMFTCKKNMEGKIVQPEN--LEYT	138
DENV3	DSRCPTQGEAILPPEEQDQNYVCKHTYVDRGWNGCGFLFGKGSVTCAKFQCLESIEGKIVQHEN--LKYT	138
WNV	KAACPTMGEAHNDKRAPFVCRQGVVDRGWNGCGFLFGKGSIDTCAKFACTKAIGRTILKEN--IKYE	138
JEV	VARCPTTGEAHNEKRADSSVCKQGFTDRGWNGCGFLFGKGSIDTCAKFSCTSKAIGRTIQPEN--IKYE	138
TBEV	AARCPMTGPATLAEHQGGTVCKRDQSDRGWGNHCGFLFGKGSIVACVKAACEAKKATGHVYDANKIVYT	140
.....		
DENV3	VVITPHSGEEHVAVNDTSEK-----GKEVKITPQSS-I-TEAELTGSGTVMTECSPRTGLDFNEMVLLQM	201
DENV3	VIIITVHTGDCQGVGNETQ-----GVTAETISQAS-T-AEAILPEYGTGLGLECSPRTGLDFNEMILLTM	199
WNV	VAIFVHGPTTVESHGNYSTQVGTQAGRFSITPAAP-S-YTLKLGEYGEVTVDCPRSGIDTNAYYVMTV	206
JEV	VGI FVHG-TTSENHGNYSQVGA-SQAAKFTVTPNAPSITLKLGDYGEVTLDCPRSGLNTAEAFVMTV	206
TBEV	VKVEPHTGDIYVAANETHSG-----RKTASFTISSE-K-TILTMGEYGDVSLLCRVASGVDLAQTIVILEL	202
=====		
DENV2	KDKAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLVTFKNPHAKKQDVVVLGSGEQEGAMHTALTGATEIQ	271
DENV3	KDKAWMVHRQWFFDLPLPWTSGATTKTPTWNRKELLVTFKNAHAKKQEVVVLGSGEQEGAMHTALTGATEIQ	269
WNV	GTKTFLVHREWFMDLNLWSSAGSTVWRNRETLMFEFEPHATKQSVIALGSGEGALHQAALAGAI PVEFSS	276
JEV	GSKSFLVHREWFHDLALPWTTPPSSTAWNRRELLMEFEEAHATQSVVALGSGEGGLHQAALAGAI VVEYSS	276
TBEV	DKTVEHLPTAQVHRDWFNDLALPWKHEGAQNWNNAERLVEFGAPHAVKMDVYNLGDQTVLLKALAGVP	272
.....		
DENV2	MSSGNLLFTGH---LKCRRLMDKLQLGMSYSMCTG-KFKVVKEIAETQHGTIVIRVQYEGDGSPCKIP	336
DENV3	TSGGTSIFAGH---LKCRRLMDKLKLKMSYAMCLN-TFVLKKEVSETQHGTILIKVEYKGEDAPCKIP	334
WNV	NTVKLTSGH---LKCRVKMEKLQLKGTTYGVCSK-AFKFLGTPADTGHGTVVLELQYTGTDGPGCKVP	339
JEV	SVKLTSGH---LKCRRLMDKLALKGTTYGMCTE-KFSFAKNPADTGHGTVVIELSYSGSDGPGCKIP	338
TBEV	VAHIBGTKYHLKSGHVTCEVGLEKLKMKGLTYTMCDDTKFTWKRAPTDSGHDTVMMEVTFSGT-KPCRIP	341
=====		
DENV2	FEIMDLEKRH-VLGRLLITVNFIVTEKDSPV--NIEAEPFGDSYIIIGVEPGQLKLNWFKK	394
DENV3	FSTEDGQGKA-HNGRLITANFVVTKEEFPV--NIEAEPFGESNIVIGIDKALKINWYRK	392
WNV	ISSVASLNDLTPVGRLLTVNPFVSVATANAKVILIELEPPFGDSYIIVGRGEQQINHHWHKS	400
JEV	IVSVASLNDMTPVGRLLTVNPFVATSSANSKVLVEMEPPFGDSYIIVGRGDKQINHHWHKA	399
TBEV	VRAVAHSGPDVNVAMLITPNPTIENNGGG--FIEMQLPPGDNIIVVG---ELSHQWFQK	395
=====		

**Fig. 5** Structure-based alignment of E protein sequences from different flaviviruses, obtained using PROCOR [49], highlighting site 1 residues in each structure using PASS (DENV2, green;

DENV3, blue; TBEV, orange; WNV, purple; JEV, dark green). The structurally equivalent residues (within 3.5 Å root mean square distance) are marked with '='



**Fig. 6** Comparison of the conformations of R1 docked to the E protein of flaviviruses. R1 is coloured according to the E protein: *green* for DENV2 (PDB ID: 1OAN); *magenta* for DENV3 (PDB ID: 1UZG); *cyan* for TBEV (PDB ID: 1SBV); *yellow* for WNV (PDB ID: 2HG0); and *pink* for JEV. The structure of DENV2 is shown in semi-transparent surface representation, with the domains coloured as in Fig. 1



has a smaller number of close contacts with DENV3 E protein, as compared to DENV2 and TBEV E proteins. The orientations of R1 docked to the E proteins from DENV2 and DENV3 are similar, but differ sufficiently to explain the difference in scores (Fig. 5). Similarly, the orientation of R1 in WNV E protein is entirely different when compared to other viruses. Multiple dockings (50 dockings in GOLD in standard default settings) of R1 to the flaviviral E proteins show that the frequency with which R1 interacts with the conserved residues of DENV2 and DENV3 are higher when compared to other flaviviral E protein interactions (data not shown). The H–N–C=O motif represents the region where maximum variation can be seen between the docked conformations. However, replacing this motif with 290 other fragments corresponding to chemical groups such as alcohols, aliphatics, esters, ethers, benzenes, amides, imides and carbocycles, did not yield compounds with improved scores, suggesting this motif is important for binding (data not shown). Modifying other groups, such as the benzotriazole and benzene groups in R1, on the other hand, could lead to better compounds that could potentially be more effective (Fig. 6).

## Conclusion

We have identified a novel inhibitor of DENV predicted to target a novel site in the E protein and interfere with the fusion process. The mode of action of this compound needs to be confirmed through E protein binding assays, fusion activity assays, mutagenesis and structural studies; at this

stage we cannot exclude the possibility that the compound targets a different site or a different protein and interferes with the viral life cycle in a different way. Improved inhibitors could be designed by substituting specific groups in the identified compound, as well as by targeting other sites such as “site 2” identified in this study, and this is the basis of ongoing investigations.

**Acknowledgments** We thank Drs Kolaskar and Kulkarni-Kale, University of Pune, India, for providing the coordinates of the structure of Japanese encephalitis virus E protein, and Charlie Huang for helpful discussions. This work was funded by a grant from the National Health and Medical Research Council (NHMRC, Australia) to PRY and BK. BK is an Australian Research Council (ARC) Federation Fellow and NHMRC Honorary Research Fellow. RY was a recipient of the 2006 Endeavour Australia Cheung Kong Award for Asian Scholars and was supported by a Research Fellowship from the Indian Council of Medical Research.

## References

- Jacobs MG, Young PR (1998) *Curr Opin Infect Dis* 11:319
- Anderson R, King AD, Innis BL (1992) *J Gen Virol* 73:2155. doi: [10.1099/0022-1317-73-8-2155](https://doi.org/10.1099/0022-1317-73-8-2155)
- Guirakhoo F, Heinz FX, Mandl CW et al (1991) *J Gen Virol* 72:1323. doi: [10.1099/0022-1317-72-6-1323](https://doi.org/10.1099/0022-1317-72-6-1323)
- Allison SL, Schalich J, Stiasny K et al (1995) *J Virol* 69:695
- Modis Y, Ogata S, Clements D et al (2004) *Nature* 427:313. doi: [10.1038/nature02165](https://doi.org/10.1038/nature02165)
- Kampmann T, Mueller DS, Mark AE et al (2006) *Structure* 14:1481. doi: [10.1016/j.str.2006.07.011](https://doi.org/10.1016/j.str.2006.07.011)
- Heinz FX, Allison SL (2001) *Curr Opin Microbiol* 4:450. doi: [10.1016/S1369-5274\(00\)00234-4](https://doi.org/10.1016/S1369-5274(00)00234-4)
- Modis Y, Ogata S, Clements D et al (2003) *Proc Natl Acad Sci USA* 100:6986. doi: [10.1073/pnas.0832193100](https://doi.org/10.1073/pnas.0832193100)



9. Rey FA, Heinz FX, Mandl C et al (1995) *Nature* 375:291. doi: [10.1038/375291a0](https://doi.org/10.1038/375291a0)
10. Bressanelli S, Stiasny K, Allison SL et al (2004) *EMBO J* 23:728. doi: [10.1038/sj.emboj.7600064](https://doi.org/10.1038/sj.emboj.7600064)
11. Morrey JD, Smee DF, Sidwell RW et al (2002) *Antiviral Res* 55:107. doi: [10.1016/S0166-3542\(02\)00013-X](https://doi.org/10.1016/S0166-3542(02)00013-X)
12. Puig-Basagoiti F, Tilgner M, Forshey BM et al (2006) *Antimicrob Agents Chemother* 50:1320. doi: [10.1128/AAC.50.4.1320-1329.2006](https://doi.org/10.1128/AAC.50.4.1320-1329.2006)
13. Zhang N, Chen HM, Koch V et al (2003) *J Med Chem* 46:4776. doi: [10.1021/jm030277k](https://doi.org/10.1021/jm030277k)
14. Knox JE, Ma NL, Yin Z et al (2006) *J Med Chem* 49:6585. doi: [10.1021/jm0607606](https://doi.org/10.1021/jm0607606)
15. Leung D, Schroder K, White H et al (2001) *J Biol Chem* 276:45762. doi: [10.1074/jbc.M107360200](https://doi.org/10.1074/jbc.M107360200)
16. Whitby K, Pierson TC, Geiss B et al (2005) *J Virol* 79:8698. doi: [10.1128/JVI.79.14.8698-8706.2005](https://doi.org/10.1128/JVI.79.14.8698-8706.2005)
17. Courageot MP, Frenkiel MP, Dos Santos CD et al (2000) *J Virol* 74:564
18. Wu SF, Lee CJ, Liao CL et al (2002) *J Virol* 76:3596. doi: [10.1128/JVI.76.8.3596-3604.2002](https://doi.org/10.1128/JVI.76.8.3596-3604.2002)
19. Chu JJ, Yang PL (2007) *Proc Natl Acad Sci USA* 104:3520. doi: [10.1073/pnas.0611681104](https://doi.org/10.1073/pnas.0611681104)
20. Mathews JH, Roehrig JT (1984) *J Immunol* 132:1533
21. Kimura-Kuroda J, Yasui K (1988) *J Immunol* 141:3606
22. Brandriss MW, Schlesinger JJ, Walsh EE et al (1986) *J Gen Virol* 67(Pt 2):229. doi: [10.1099/0022-1317-67-2-229](https://doi.org/10.1099/0022-1317-67-2-229)
23. Shimoni Z, Niven MJ, Pitlick S et al (2001) *Emerg Infect Dis* 7:759
24. Liao M, Kielian M (2005) *J Cell Biol* 171:111. doi: [10.1083/jcb.200507075](https://doi.org/10.1083/jcb.200507075)
25. McCown M, Diamond MS, Pekosz A (2003) *Virology* 313:514. doi: [10.1016/S0042-6822\(03\)00341-6](https://doi.org/10.1016/S0042-6822(03)00341-6)
26. Bai F, Wang T, Pal U et al (2005) *J Infect Dis* 191:1148. doi: [10.1086/428507](https://doi.org/10.1086/428507)
27. Adelman ZN, Sanchez-Vargas I, Travanty EA et al (2002) *J Virol* 76:12925. doi: [10.1128/JVI.76.24.12925-12933.2002](https://doi.org/10.1128/JVI.76.24.12925-12933.2002)
28. Bai F, Town T, Pradhan D et al (2007) *J Virol* 81:2047. doi: [10.1128/JVI.01840-06](https://doi.org/10.1128/JVI.01840-06)
29. Hrobowski YM, Garry RF, Michael SF (2005) *Virol J* 2:49. doi: [10.1186/1743-422X-2-49](https://doi.org/10.1186/1743-422X-2-49)
30. Marks RM, Lu H, Sundaresan R et al (2001) *J Med Chem* 44:2178. doi: [10.1021/jm000412i](https://doi.org/10.1021/jm000412i)
31. McInnes C (2007) *Curr Opin Chem Biol* 11:494. doi: [10.1016/j.cbpa.2007.08.033](https://doi.org/10.1016/j.cbpa.2007.08.033)
32. Kirchmair J, Distinto S, Schuster D et al (2008) *Curr Med Chem* 15:2040. doi: [10.2174/092986708785132843](https://doi.org/10.2174/092986708785132843)
33. Coupez B, Lewis RA (2006) *Curr Med Chem* 13:2995. doi: [10.2174/092986706778521797](https://doi.org/10.2174/092986706778521797)
34. Yang JM, Chen YF, Tu YY et al (2007) *PLoS ONE* 2:e428. doi: [10.1371/journal.pone.0000428](https://doi.org/10.1371/journal.pone.0000428)
35. Brady GP Jr, Stouten PF (2000) *J Comput Aided Mol Des* 14:383. doi: [10.1023/A:1008124202956](https://doi.org/10.1023/A:1008124202956)
36. Landau M, Mayrose I, Rosenberg Y et al (2005) *Nucleic Acids Res* 33:W299. doi: [10.1093/nar/gki370](https://doi.org/10.1093/nar/gki370)
37. Zhang W, Chipman PR, Corver J et al (2003) *Nat Struct Biol* 10:907. doi: [10.1038/nsb990](https://doi.org/10.1038/nsb990)
38. Bodian DL, Yamasaki RB, Buswell RL et al (1993) *Biochemistry* 32:2967. doi: [10.1021/bi00063a007](https://doi.org/10.1021/bi00063a007)
39. Lipinski CA, Lombardo F, Dominy BW et al (1997) *Adv Drug Deliv Rev* 23:3. doi: [10.1016/S0169-409X\(96\)00423-1](https://doi.org/10.1016/S0169-409X(96)00423-1)
40. Verdonk ML, Cole JC, Hartshorn MJ et al (2003) *Proteins* 52:609. doi: [10.1002/prot.10465](https://doi.org/10.1002/prot.10465)
41. Rarey M, Kramer B, Lengauer T et al (1996) *J Mol Biol* 261:470. doi: [10.1006/jmbi.1996.0477](https://doi.org/10.1006/jmbi.1996.0477)
42. Kuntz ID, Blaney JM, Oatley SJ et al (1982) *J Mol Biol* 161:269. doi: [10.1016/0022-2836\(82\)90153-X](https://doi.org/10.1016/0022-2836(82)90153-X)
43. Wang R, Lai L, Wang S (2002) *J Comput Aided Mol Des* 16:11. doi: [10.1023/A:1016357811882](https://doi.org/10.1023/A:1016357811882)
44. Kontoyianni M, McClellan LM, Sokol GS (2004) *J Med Chem* 47:558. doi: [10.1021/jm0302997](https://doi.org/10.1021/jm0302997)
45. Wallace AC, Laskowski RA, Thornton JM (1995) *Protein Eng* 8:127. doi: [10.1093/protein/8.2.127](https://doi.org/10.1093/protein/8.2.127)
46. Nybakken GE, Nelson CA, Chen BR et al (2006) *J Virol* 80:11467. doi: [10.1128/JVI.01125-06](https://doi.org/10.1128/JVI.01125-06)
47. Kolaskar AS, Kulkarni-Kale U (1999) *Virology* 261:31. doi: [10.1006/viro.1999.9859](https://doi.org/10.1006/viro.1999.9859)
48. Modis Y, Ogata S, Clements D et al (2005) *J Virol* 79:1223. doi: [10.1128/JVI.79.2.1223-1231.2005](https://doi.org/10.1128/JVI.79.2.1223-1231.2005)
49. Subbarao N, Haneef I (1991) *Protein Eng* 4:877. doi: [10.1093/protein/4.8.877](https://doi.org/10.1093/protein/4.8.877)