Available online at www.sciencedirect.com
SCIENCE DIRECT
Bioorganic & Medicinal Chemistry 14 (2006) 2674–2683
Bioorganic & Medicinal Chemistry

A modelling study of a non-concerted hydrolytic cycloaddition reaction by the catalytic antibody H11

Rachel L. Clark,[a] Blair F. Johnston,[a] Colin J. Suckling[b] and Simon P. Mackay[a,*]

[a]*Department of Pharmaceutical Sciences, University of Strathclyde, Glasgow G4 0NR, UK*
[b]*Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow G1 1XL, UK*

Received 12 July 2005; revised 23 November 2005; accepted 28 November 2005
Available online 27 December 2005

Abstract—H11 is the first antibody reported to have dual activity as a non-concerted, Diels–Alderase and hydrolytic catalyst. It was previously shown to catalyse the cycloaddition of acetoxybutadiene **1a** to *N*-alkyl maleimides **2** to afford hydroxy-substituted bicyclic adducts **3** with a 30% ee of a major isomer. To better understand this mechanism and the partial stereospecificity, a homology model of H11 was constructed and used in docking studies to evaluate potential antibody–ligand complexes. The model suggested the hydrolytic nature of H11 was due to Glu 95H acting as a catalytic base, and evaluation of the shape complementarity of the proposed antibody–ligand complexes confirmed at a semi-quantitative level the observation that the major enantiomer is produced in a 30% ee.

1. Introduction

The ability of antibodies to catalyse chemical reactions of synthetic interest has been widely studied[1] and several reaction types have been successfully demonstrated including aldol reactions,[2] epoxidation[3] and hydrolytic resolution.[4] Cyclisation reactions featured also including peptide cyclisation,[5] carbocation[6] and electrocyclic reactions.[7] In our laboratories, an antibody designated H11 was identified[8] and shown to catalyse the cycloaddition of acetoxybutadiene **1a** to *N*-alkyl maleimides **2** to afford hydroxy-substituted bicyclic adducts **3** (Fig. 1). Investigation of the mechanism of action of this antibody showed that the primary catalytic event was probably the hydrolysis of acetoxybutadiene **1a** by the antibody affording the corresponding ene–diol **1b**, which rapidly reacted with maleimide **2** at the binding site of the antibody.[9] Untrapped ene–diol **1b** appeared as crotonaldehyde **4**. This is the first instance of an antibody catalysing both a hydrolytic and cycloaddition reaction type within the same binding site, which makes the molecular events of the reaction mechanism of H11 an interesting case study. Through chemical modification, we have demonstrated that the acidic amino acids in

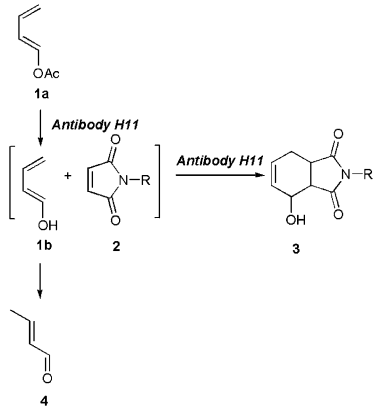

Figure 1. Reactions catalysed by H11.

association with tyrosine and histidine appear to be important in these reactions, whilst a role for tryptophan was inferred from fluorescence measurements.[9] Finally, a significant degree of stereoselectivity was shown in the cycloaddition by NMR analysis of Mosher's ester derivatives of the hydroxy adduct.[10] This reaction thus provides for the preparation of a polyfunctional, bicyclic compound with three defined stereocentres (Fig. 2).

Although preparative quantities of product could be isolated, the stereoselectivity of the reaction was not

Keywords: Modelling; Cycloaddition; Hydrolysis; Antibody; Diels-Alderase.
* Corresponding author. Tel.: +44 141 548 2866; fax: +44 141 552 6443; e-mail: simon.mackay@strath.ac.uk

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.11.042

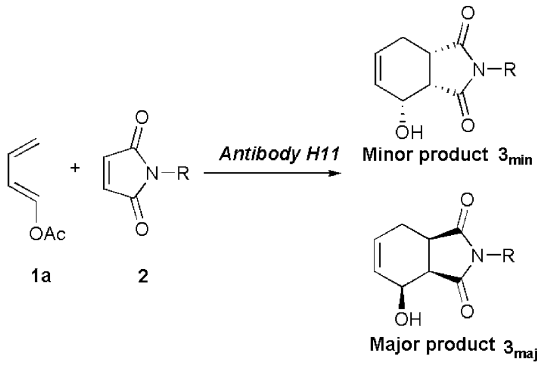

Figure 2. Stereoselective preparation of polyfunctional, bicyclic compounds **3**.

sufficient to make H11 of use in synthesis. In order to better understand the molecular reaction mechanism, the amino acid sequence for H11 was obtained.[11] Since the structures of antibodies are very regular outside of the variable loops, a useful model of an antibody can be obtained by comparative modelling, an approach that has been used successfully with catalytic antibodies.[12] With the sequence information, the chemical modification data, and the stereochemical course defined, there was a sufficient basis to create a model of H11 from which the requirements for an improved synthetic antibody could be deduced. We now report, how we used several parallel computational procedures to construct such a model of H11 that accounts for the observed properties of the antibody.

2. Homology modelling

Suitable template proteins for the variable heavy (V_H) and variable light (V_L) fragments of H11 were obtained by a BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) search of the Protein Databank (PDB). The V_H fragment was found to have 97.4% identity with anti-hen egg white lysozyme antibody D1.3 (PDB entry 1A2Y[13]) of which all three complementary-determining regions (CDRs) were identical. The BLAST search of the PDB with the amino acid sequence of the V_L fragment revealed high identity (81.7%) with the V_L of a human antibody; M4L/Y(27D)D/T94H Mutant of Len (PDB entry 1EEQ[14]). The CDRs of H11 and the template 1EEQ share 85.3% identity (Fig. 3). The light chain template antibody has two residue substitutions within CDRL1 compared to H11; residues 27E and 30; aspartic acid and serine, respectively, in 1EEQ are substituted for tyrosine and glutamine in H11. Although the substituted amino acids do not share any similarity, the residues are not in the vicinity of the proposed active site and were not shown to have a significant effect on the overall structure, assessed through protein health validation (described in Section 8.1). Structural differences were noticed within CDRL3, where the amino acid sequence of H11 was found to have a PY deletion (residues 96 and 97), although the proline in question does not invoke a significant change in direction of the backbone of the template V_L. Moreover, this deletion had little effect on the overall structure of the backbone of H11 compared to 1EEQ, and the observation was made that this deletion made the active site cleft larger than the equivalent in 1EEQ. Two further residues in CDRL3 are substituted for similar residues in H11, namely Gln91 and Ser99 in the template V_L are replaced by His91 and Thr99 in H11 V_L. These substitutions had no detrimental effect on the validity of the homology model 3D structure. Homology models of each variable fragment were constructed using these templates as described in Section 8.

3. Macromolecular protein–protein docking

The V_H chain of H11 had such high sequence identity to the template antibody that the assumption was made that the V_L fragment of H11 would interact with the V_H fragment in a similar way to the variable fragments of the lysozyme antibody. Where this approach has been adopted in previous studies,[12b,15] the independently generated V_H and V_L models have been superimposed onto the backbone atoms of the template with which the highest sequence identity was found. As an alternative method to superimposition, we docked the V_H and V_L fragments of H11 using the protein docking and molecular superposition software Hex 4.2 (http://www.csd.abdn.ac.uk/hex/), recently evaluated by the critical assessment of predicted interactions (CAPRI), a community-wide experiment to assess the quality of protein docking methods to predict protein–protein interactions.[16,17] The software uses polar Fourier correlations in order to rapidly locate potential protein–protein complexes by the rotational search of one protein with the other.[18] The quality of suggested complex poses is then quantified in terms of their shape complementarity and electrostatic interactions (E_{shape} and $E_{electrostatic}$, respectively).[19] Potentially, millions of docking conformations can be generated with several thousand of the best results (most negative energy), then treated with a rigid-body energy minimisation algorithm to further

CDRL1

	25	26	27	27A	27B	27C	27D	27E	27F	28	29	30	31	32	33	34	35
H11	K	S	S	Q	S	V	L	Y	S	S	N	Q	K	N	Y	L	A
1EEQ	K	S	S	Q	S	V	L	D	S	S	N	S	K	N	Y	L	A

CDRL2

	51	52	53	54	55	56	57
H11	W	A	S	T	R	E	S
1EEQ	W	A	S	T	R	E	S

CDRL3

	91	92	93	94	95	96	97	98	99	100	101
H11	H	Q	Y	L	S	H	-	-	T	F	G
1EEQ	Q	Q	Y	Y	S	H	P	Y	S	F	G

Figure 3. The alignment of the CDR regions of H11 V_L with the template 1EEQ, identical residues highlighted red, similar residues highlighted green, numbered according to H11 V_L using the Kabat system (http://www.bioinf.org.uk/abs/chothia.html).

optimise the pose. The ten lowest energy poses identified are shown in Table 1.

Based on some general properties of antibodies and more specific experimental observations with H11, we used the additional criteria below to filter and assess the validity of the ten docked fragments and identified the fourth pose as meeting all the requirements (Fig. 4).

- The backbone of the docked fragments should resemble the backbone of the class of antibody structures that 1A2Y is representative of.
- The loop regions should be at the interface of the two chains, on the same side of the molecule where they would be in direct contact with the antigen.[20]

Table 1. The ten lowest energy poses found by Hex (kJ mol^{-1})

Pose	E_{total}	E_{shape}	$E_{electrostatic}$
1	−538.08	−490.46	−47.63
2	−533.14	−501.41	−31.73
3	−531.81	−462.61	−69.20
4	−529.57	−502.99	−26.58
5	−525.82	−518.60	−7.22
6	−525.42	−502.64	−22.78
7	−515.16	−474.84	−40.32
8	−505.45	−400.16	−105.29
9	−504.50	−493.97	−10.53
10	−501.55	−476.04	−25.51

- The distance between the *C*-terminal of the light chain and the *N*-terminal of the heavy chain should be 35–40 Å (the distance covered by an engineered bridging sequence $(GGGGS)_3$ that was spliced to the terminal residues to ensure the antibody did not dissociate at the low concentrations employed for cycloaddition reactions).[11]
- Where the distance between the N and C terminus was less than 40 Å, the bridging sequence was constructed between the two termini and subjected to a brief dynamics simulation with the chains fixed. If the linker occupied the looped regions, the structure was rejected based on its proposed interference with hapten binding.

When we superimposed the independently generated V_H and V_L models onto the backbone atoms of the template structure 1A2Y with which H11 V_H had 97.4% sequence identity using the more common approach of antibody homology modelling,[12b,15] the RMSD (all atoms) with the Hex generated complex was 2.34 Å (Fig. 5), which further confirms Hex as a useful programme in the armoury of tools available for modelling protein–protein interactions.

4. Identification of the active site

The docked heavy and light chains identified as the likely complex for the H11 antibody were minimised

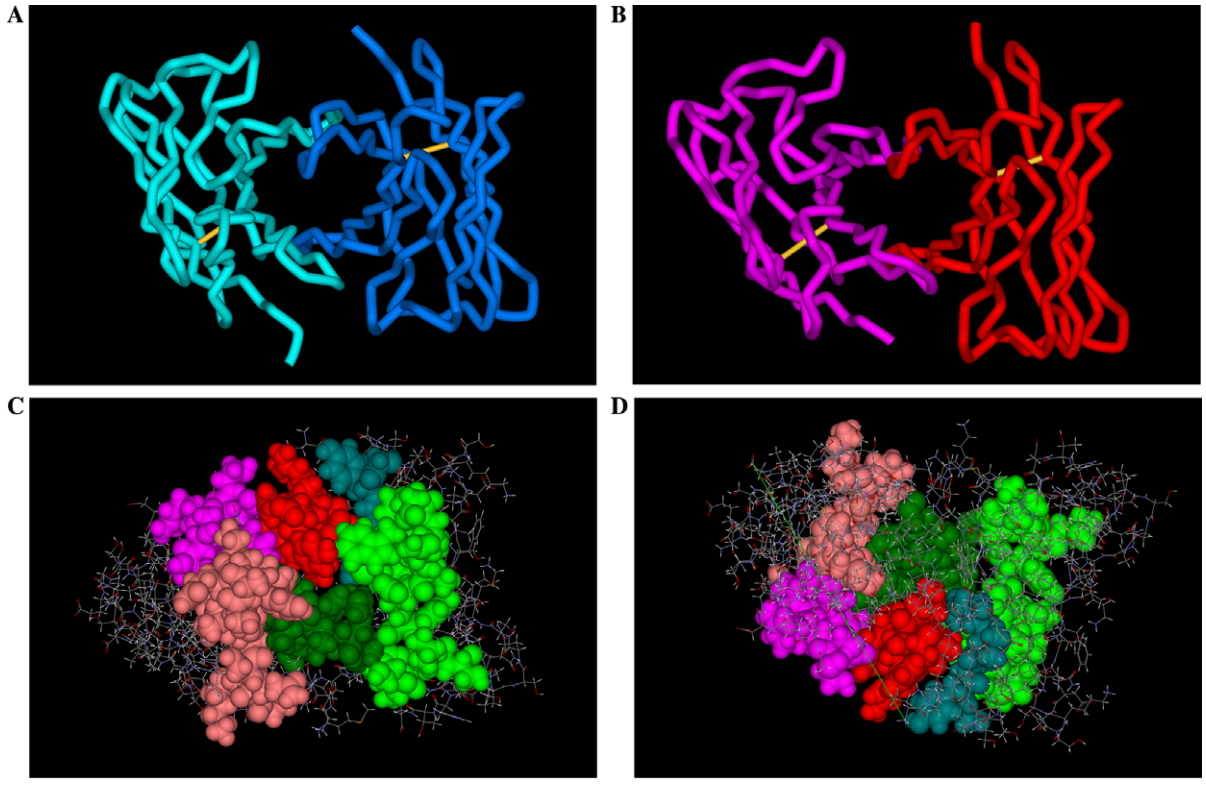

Figure 4. The backbone of the docked heavy and light chains (pose 4) using the (A) Hex software (dark blue and light blue, respectively) and the (B) backbone of the antibody complex crystal structure used to generate the model heavy chain (1A2Y) (heavy chain, red; light chain pink). (C) The looped regions of the H11 model in CPK representation (green shades; heavy chain, pink shades; light chain). (D) The distance between the N and C terminal residues of the two chains (35.03 Å) where the bridging peptide links.

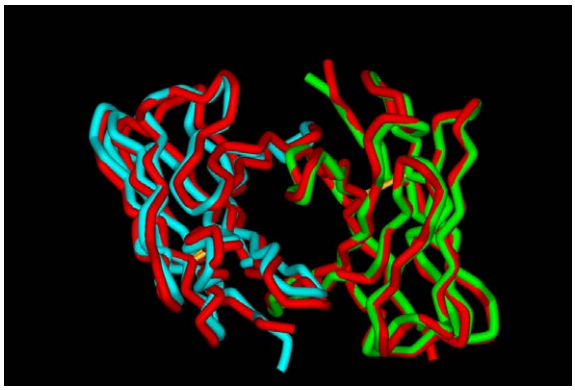

Figure 5. The backbone of the Hex docked heavy and light chains (red) of H11 superimposed with the docked heavy (green) and light chain (blue) using the superimposition method.

using the standard step-wise restraint protocol for proteins. In order to identify potential binding sites in H11 where catalysis could take place, the LigandFit software,[21] which uses a flood filling algorithm to identify crevices on the surfaces of proteins, located a hydrophobic pocket that contained Glu 95H (Fig. 6A) and

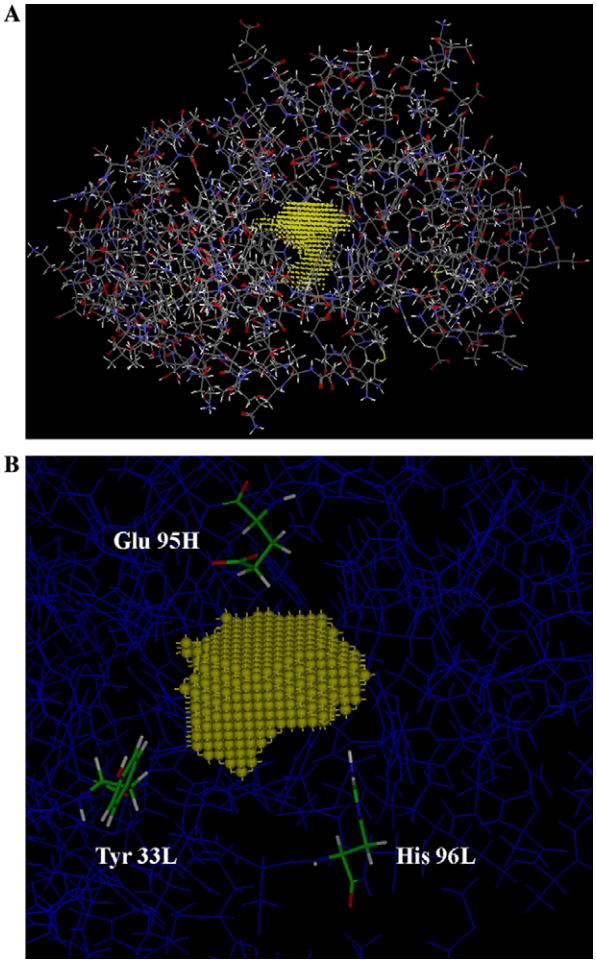

Figure 6. (A) The binding pocket suggested by LigandFit's flood filling algorithm (shown in yellow). (B) The results from LigandFit highlighting key residues in the active site designated H or L for the heavy and light chains, respectively.

was the first and largest cleft identified. Seven sites were identified in total; the second site was of comparable size, but was buried deep in the protein, away from the loop region and the remaining sites were much smaller and only made contacts with residues from one of the variable chains.

Our earlier experimental studies had shown that an acidic residue in the looped region of H11 was involved in the catalytic activity of the antibody.[8,9] Apart from Glu 95H, only two other acidic residues were in the CDRs of our model, and both were on the surface of the protein. Moreover, the cleft that contained Glu 95H (CDRH3) was also in close proximity to other residues identified as having a role in catalysis,[9] namely a tyrosine (Tyr 33L, CDRL1) and a histidine (His 96L, CDRL3), confirming that this was the likely catalytic site (Fig. 6B).

The role of a glutamic acid residue as a general base that abstracts a proton from a nearby water molecule has been proposed in several reaction mechanisms, including: TEM-1, a β-lactamase enzyme that catalyses the hydrolysis of the lactam bond;[22,23] a phosphotransfer reaction catalysed by pyruvate dehydrogenase kinase[24] and peptide hydrolysis catalysed by a leucine aminopeptidase enzyme of *Aeromonas proteolytica*.[25] Glutamate is also the key catalytic residue in lysozyme enzymes, which hydrolyse glycosidic bonds. Several catalytic antibodies are also reported to have a glutamate or an aspartate residue acting as a general base during catalysis;[1d] antibody 34E4, which catalyses the conversion of benzisoxazoles to salicylonitriles, relies on a key interaction between Glu 50H and the hapten for the optimum rate of catalysis;[26] antibody 4B2 catalyses the allylic isomerisation of β,γ-unsaturated ketones involving the stabilisation of an enol intermediate through Glu 34L and Asp95 residues in the active site[27,28] and the catalytic antibody 14D9 promotes the enantioselective protonation of prochiral enol ethers by the formation of a network of hydrogen bonds via the side chains of Asp 101H, Tyr 36L and a water molecule within the active site.[29]

For a glutamate to act as a general base in catalysis, it tends to be within a protein cleft rather than on the surface. This fact, coupled with the high sequence identity of the V_H chain of H11 with the V_H chain of the template lysozyme antibody, suggests a similar catalytic mechanism is responsible for the activity of H11 that could well involve Glu 95H.

Hydrolysis of the acetoxybutadiene to the corresponding ene–diol prior to cycloaddition to the bicyclic product was shown to proceed most rapidly at pH 7–8.[8] In common with other antibody mediated hydrolytic procedures,[22,25] at this pH, Glu 95 is potentially in an ionisation state that could involve proton abstraction from an in situ molecule of water as part of the catalytic mechanism. Examination of the crystal structure of the template lysozyme antibody used to model the H11 V_H fragment (1A2Y) revealed a water molecule co-crystallised 2 Å from the corresponding glutamate, which

was transferred to the model of H11 by superimposition of the two V_H fragments according to sequence alignment (RMSD 0.74 Å). Based on our observation that hydrolysis of the alkoxy diene promotes cycloaddition[10] and that hydrogen bonding can play a key role both electronically and structurally in such a mechanism, we proposed the series of events shown in Figure 7, where the hydrogen bond donor roles could be fulfilled by either tyrosine or histidine residues.

5. Docking studies

The H11 antibody was raised to the bovine serum albumin conjugate hapten **5**[8] and subsequent investigation of its capability revealed an ability to catalyse the reactions shown in Figure 2.[10] Kinetic measurements established that the *N*-alkyl substituent of the maleimide did not play a major role in binding because the K_M values for the *N*-butyl and *N*-ethyl derivatives were of the same order.[10] In contrast, the K_M values for the acetoxybutadiene imply its binding with H11 is a major determinant of activity.[9] Given that shape complementarity is a feature commonly observed in Diels–Alderase antibodies that is emphasised by the importance of the relative orientation of the reagents (which is reflected in the adduct topography[30]), we performed our initial docking studies using **5a$_{maj}$**, which is a simplified analogue of the hapten to which H11 was originally raised.

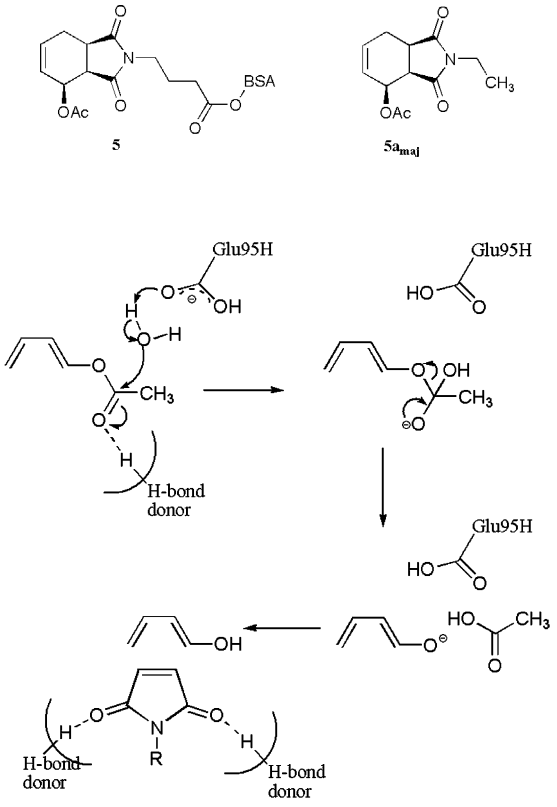

Figure 7. The proposed catalytic mechanism for the hydrolysis of acetoxybutadiene. The H-bond donor in the hydrolytic step serves to render the acetoxy carbonyl group more susceptible to nucleophilic attack by the deprotonated water molecule. The H-bond donors in the cycloaddition step play a structural role in orientating the incoming maleimide dienophile with the diene.

The ligand **5a$_{maj}$** was docked manually into the binding site identified by the LigandFit algorithm, a region that showed significant shape complementarity between the hapten and the antibody even before further refinements were performed (Fig. 8).

The manually docked ligand was then subjected to simulated annealing treatment of the binding site within the structure, and a subsequent minimisation. Residues within 8 Å of the substrate were included in the structural subset to be annealed. This procedure was repeated for **5a$_{min}$** and the two proposed orientations are shown in Figure 9.

Two hydrogen bonds were observed in the complex with **5a$_{maj}$**, between the carbonyl oxygen atoms of the maleimide and both Tyr 33L and His 96L as donors. However, the complex with the adduct with opposite stereochemistry **5a$_{min}$**, formed no hydrogen bonds with the protein, and moreover, was not bound as deeply in the cleft as **5a$_{maj}$**, which prevented hydrogen bond formation with the same atoms of the protein. In both cases, the acetoxy group is positioned in the vicinity of the proposed catalytic residue Glu 95H, with the methyl group occupying a hydrophobic cleft bordered by Val 37H, Trp 47H, Met 50H and Phe 100L. The *N*-alkyl substituent, which does not play a major role in binding,[10] lies in a large exit channel from the catalytic site where the hydrogen bonding Tyr 33L and His 96L residues make up the walls and Leu 94L occupies the floor.

Based on the orientations of the adduct with respect to the binding site in the complexes, it would appear that the diene can approach the antibody in two alternative orientations (Fig. 10), which, if adopted equally, would result in a non-stereospecific production of **3**.

The dienes **1a$_{maj}$** and **1a$_{min}$** were each treated with the flexible docking and minimisation procedure previously described for adducts **5a**. Surprisingly, the resulting diene–antibody complexes showed there was only a favourable hydrogen bond formed between the diene and the protein when docked in the minor orientation. Clearly, whilst hydrogen bonding is instrumental in

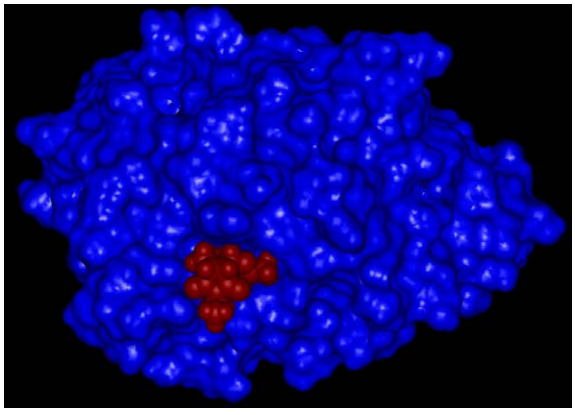

Figure 8. The shape complementarity observed in the homology model on initial docking of substrates.

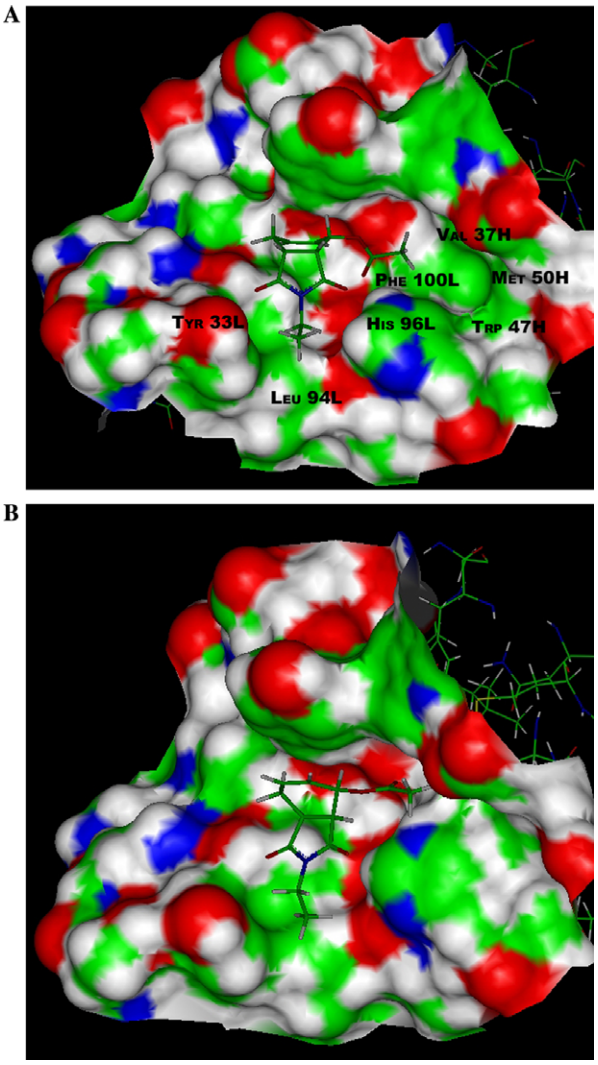

Figure 9. Proposed alternative docking orientations that would results in product formation with opposite chirality (A) major isomer (B) minor isomer.

catalysis, overall binding energies will be the major determinant of activity. To quantify the overall stability of the complexes, the binding enthalpy (Table 2) and the accessible surface area (Table 4) were calculated for the adducts **5a** and the dienes **1a**. Binding enthalpy was computed by summing the potential energies of the free ligand and receptor after separation from the complex followed by minimisation, then subtracting this value from the potential energy of the complex itself. Such a calculation provides information about the non-bonded contribution to complexation and the internal strains on both ligand and receptor induced through the binding event.

Table 2. Computed enthalpy terms (kcal mol^{-1}) for the ligand–antibody complexes

Ligand	E_{bind}[a]	$E_{complex}$	E_{ligand}	$E_{receptor}$
5a$_{maj}$	−24.39	−2035.08	9.20	−2019.89
5a$_{min}$	−20.64	−2040.52	6.48	−2026.36
1a$_{maj}$	−13.59	−2021.98	17.02	−2025.41
1a$_{min}$	−12.42	−2042.16	12.69	−2042.44

[a] $E_{bind} = E_{complex} - (E_{ligand} + E_{receptor})$.

Table 3. The ratio of bound conformers and product diastereoisomers, where $R = 8.314$ J K^{-1} mol^{-1} and $T = 300$ K

Ligand	E_{bind}[a]	Boltzmann factor[b]	Population	Ratio
5a$_{maj}$	−24.39	exp (24.39/2.49)	17,947.0	82
5a$_{min}$	−20.64	exp (20.64/2.49)	3980.5	18
1a$_{maj}$	−13.59	exp (13.59/2.49)	234.6	62
1a$_{min}$	−12.42	exp (12.42/2.49)	5.0	38

[a] $E_{bind} = E_{complex} - (E_{ligand} + E_{receptor})$.
[b] Boltzmann factor $= \exp(\frac{-E}{RT})$.

Table 4. The Connolly surface area (Å^2) for the ligand–antibody complexes

Ligand	Active site	Complex	Ligand	Solvent accessible area of ligand in complex
5a$_{maj}$	1161.43	1094.03	242.72	111.88
5a$_{min}$	1190.25	1231.27	250.80	163.89
1a$_{maj}$	1155.87	1131.97	153.67	88.22
1a$_{min}$	1152.18	1125.61	155.59	90.35

The expected ratios of each diene conformation and adduct can be calculated using the equation for the Boltzmann factor, which can determine the population of each conformer and its relative ratios (Table 3).[31]

Chiral derivatisation of the products of H11 revealed that the product **3a$_{maj}$** was obtained in 30% enantiomeric

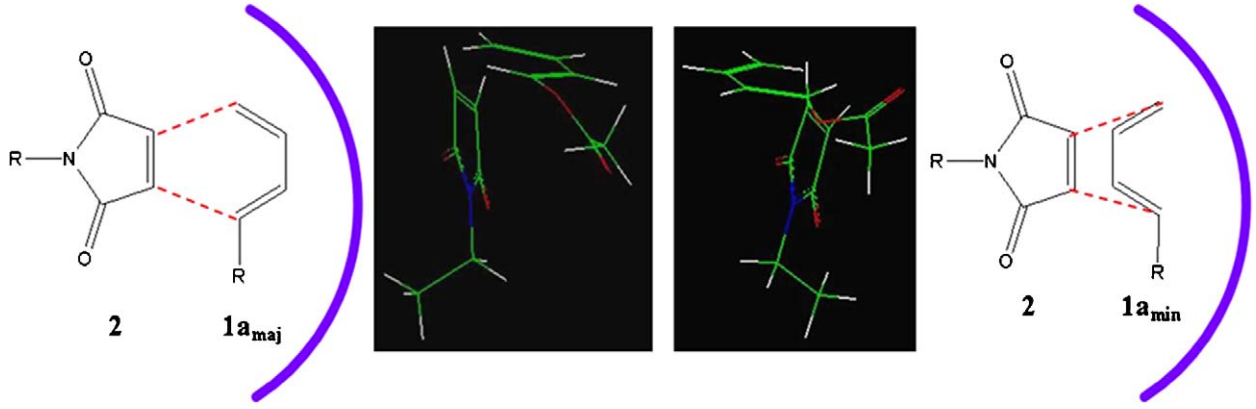

Figure 10. The alternative transition states resulting in **5** having opposite stereochemistry.

excess to the minor isomer $\mathbf{3a_{min}}$. Analysis of the relative binding enthalpies of two enantiomers of hapten **5a** and the two orientations of the diene $\mathbf{1a_{maj}}$, $\mathbf{1a_{min}}$ revealed the binding enthalpy was more favourable for the hapten and diene with the stereochemistry and orientation which forms the major product $\mathbf{3a_{maj}}$, which is in agreement with the experimental data.[10] When these computed enthalpies are treated to a Boltzmann population calculation,[31] the relative populations of the adducts broadly conform to the experimental values. Indeed, the dienes themselves, when bound to the antibody, reflect a conformation population that would produce an isomeric excess of one adduct and is also broadly in line with our experimental results.

To further investigate this stereochemical aspect of the mechanism, we calculated the Connolly accessible surface areas of the different complexes and showed that the major isomer should be preferred over the minor isomer; for each diene, the orientation resulting in the major product is able to bind deeper into the cleft, with the solvent accessible area for the $\mathbf{1a_{maj}}$ being 88.22 Å^2 and $\mathbf{1a_{min}}$ 90.35 Å^2. The smaller solvent accessible area is also seen with the $\mathbf{5a_{maj}}$ complex compared to the $\mathbf{5a_{min}}$, with the hapten bound deeper in the active site cleft. A deeper binding orientation achieves three objectives; first, hydrogen bonding with the relevant residues in the cleft is maximised; second, the hydrolysable ester group can be accessed by the glutamate/water catalytic residue; third, exposing less of a hydrophobic surface is entropically favourable in an aqueous medium.

6. Molecular dynamics simulations

In order to assess the structural integrity of our antibody model, we performed a series of dynamics simulations on selected complexes. The free antibody and the complex with the $\mathbf{5a_{maj}}$ adduct were subjected to extended constant temperature dynamics (>1 ns) using explicit water representation under periodic boundary conditions. Sampled average structures from each of the runs were minimised and superimposed to evaluate stability and movement within the protein model. Figure 11A indicates that the main chain backbone is stable and does not undergo significant movement during the extended simulation. The area of greatest conformational freedom is in the binding site region and is highlighted in Figure 11B. The histidine and tyrosine residues show a degree of flexibility in their side chains, which would be expected to allow trapping and release of reagents and products. Significantly, the loop that contains the catalytic glutamate residue appears to undergo more noticeable rearrangement, particularly between the unbound and bound forms of the antibody. Again, such a rearrangement would be expected for complexation and could suggest a key positioning dynamic between the catalytic residue and the reagent/product. This kind of loop movement during catalytic events in antibodies is not unique, particularly in hydrolytic processes; in the antibody 7F11,[15a] which catalyses the decomposition of the benzoate ester of an aryl adamantyl 1,2 dioxetane, there is significant movement observed at the interface of the heavy and light chains and the

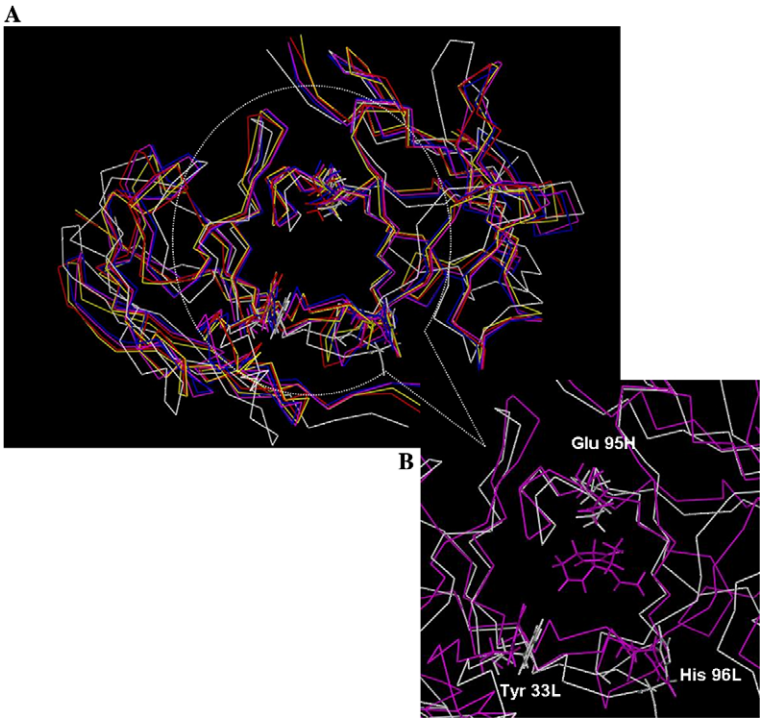

Figure 11. (A) The backbone of H11 after the free antibody (white) and the complex $\mathbf{5a_{maj}}$ (sampled average structures, red, yellow, magenta and blue) were subjected to extended constant temperature dynamics (>1 ns). The backbone atoms of the averaged complex structures were superimposed with RMS values of 0.8–2.3 Å, the backbone of the free antibody superimposed onto the complex with an RMS of 2.5 Å. (B) An expansion of the central antigen binding site at the interface of the heavy and light chains, the averaged minimised complex (magenta) containing ligand $\mathbf{5a_{maj}}$ superimposed with the free antibody (white) highlighting the movement seen in the key residues Glu 95H, Tyr 33L and His 96L.

catalytic residues in the active site. This movement is reduced in the hapten–antibody complex, which is also observed in our dynamics simulations (Fig. 11A).

7. Conclusion

The model of the antibody H11 we have built has produced a greater understanding of its dual catalytic activity as a Diels–Alderase and hydrolytic catalyst. The protein docking software Hex successfully associated the two variable fragments of the antibody in a manner that could explain catalytic activity and was confirmed by an RMS value of 2.34 Å with the model generated by the standard superimposition method. The experimental data of H11 previously reported by the group have been used to validate the model, and confirmed at a semi-quantitative level, the observation that the major enantiomer is produced in a 30% ee. The proposed mechanism of action of hydrolysis by a glutamate residue acting as a catalytic base is consistent with other hydrolytic antibodies and enzymes. The importance of shape complementarity in other Diels–Alderase antibodies is also seen in H11.

8. Experimental

8.1. Homology modelling

The 3D structure of the variable region of H11 was constructed by comparative modelling using Discovery Studio (DS) Modeling 1.1[32] and InsightII on a Hewlett Packard Workstation xw8000[33] and a Silicon Graphics Octane2 workstation[34] (2×600 MHz MIPS R14000 processors), respectively. Suitable template proteins were obtained through a BLAST search of the Protein Databank (PDB) using the BLOSUM 62 matrix with a gap penalty of 11 and a gap extension penalty of 1. The V_H fragment had 97.4% identity with anti-hen egg white lysozyme antibody D1.3 (PDB entry 1A2Y[13]) of which all three complementary-determining regions (CDRs) were identical. The BLAST search of the PDB with the amino acid sequence of the V_L fragment revealed high identity (81.7%) with the V_L of a human antibody; M4L/Y(27D)D/T94H Mutant of Len (PDB entry 1EEQ[14]). The CDRs of H11 and the template 1EEQ share 85.3% identity.

The amino acid sequences of the V_H and V_L regions of H11 were aligned to their 3D templates using the 'Align123' experiment of the software package DS Modelling 1.1. The resulting alignment was then used to build models of each variable region using the MODELER software, where three models of each variable region were built at the high level of optimisation and assessed according to their probability density function (PDF) energy. The models with the lowest PDF energy were then evaluated using the protein health validation algorithm to measure the overall compatibility of the amino acid sequence in the environment of the 3D structure.[35] The minimised structures of H11 V_H and V_L both had high scores indicating the side chain of each amino acid in the sequence is in a satisfactory environment and the protein is likely to be folded satisfactorily.

To assess the validity of the homology models, they were minimised in a step wise manner fixing the heavy atoms followed by the backbone atoms before the full structure was allowed to relax. The cff91 forcefield was used for all energy calculations, and models were minimised using a distance dependant dielectric of 4.0, with non-bonded cutoffs treated separately for van der Waals forces (9.5 Å) and electrostatic interactions (14 Å). Minimisation was effected using the Conjugate Gradients algorithm until an energy convergence criterion of 0.1 kcal mol^{-1} $Å^{-1}$ was reached. These structures were further assessed by evaluating the Ramachandran map, which showed that the V_H and V_L fragments had 94.83% and 95.69% of the residues, respectively, in the allowed region. Such a high percentage indicates that in both models the backbone dihedral angles ϕ and ψ are satisfactory.

8.2. Macromolecular docking with Hex 4.2

The minimised structures of the heavy chain (receptor) and light chain (ligand) were loaded into Hex as ·pdb files. The ligand was positioned in the vicinity of the receptor and the docking control parameters set to ensure that the light chain was fully able to explore the surface of the heavy chain. Docked poses were assessed on both shape complementarity and electrostatic interactions. Subsequently, similar poses were grouped into clusters based on a RMSD of 2.0 Å from the pose automatically chosen as the cluster centre. The ten best clustered docking poses (lowest energy) were then assessed according to the criteria listed in Section 3.

8.3. Identification of the active site

The docked light and heavy chain complex was then minimised as described previously (Section 8.1)[34] to relax any bad surface contacts introduced in the rigid-body Hex docking study. The optimised structure was then imported into Cerius[2] (Accelrys, version 4.8.1)[21] and treated with a flood fill algorithm in order to identify the active site. The first and largest cleft of the seven identified contained the proposed catalytic residue Glu 95H.

8.4. Substrate docking

Atomic coordinates for the major and minor *N*-ethyl derivatives **5a** of hapten **5** and the diene **1a** were constructed from the fragment database from the InsightII program, and were minimised using cff91 forcefield and parameters previously described. The optimised structures were docked manually into the homology model, guided by the distinct shape complementarity observed in the active site region previously identified by the flood filling algorithm (Fig. 8). Residues within 8 Å of the substrate were included in the structural subset and subjected to simulated annealing to optimise the complex from 300 to 270 K, followed by minimisation using the

parameters previously described. The enthalpies of the resulting minimised structures were computed and derived using the equation $E_{bind} = E_{complex} - (E_{ligand} + E_{receptor})$; the ligand and receptor enthalpies were each calculated separately once removed from the complex and minimised (Table 2).

8.5. Molecular dynamics

Molecular dynamics (MD) simulations were carried out on a 3.2 GHz dual Intel Xeon workstation with 4 Gb of RAM, using InsightII (version 2000.3) and the Discover molecular simulation software. Two simulations were setup; one with free H11, and one with H11 complexed with **5a$_{maj}$**. H11 was centred in a rectangular periodic cell with a side length of $80.0 \times 70.0 \times 80.0$ Å and soaked in a box of SPC water. The solvated protein was energy minimised using the Conjugate Gradients algorithm, with the van der Waals and electrostatic interactions treated separately as before until an energy convergence criterion of 1.0 kcal mol^{-1} Å^{-1} was reached. MD simulations on minimised solvated structures were performed by integrating Newton's equations of motion using the Verlat-leapfrog algorithm for 800 ps at a simulated temperature of 298 K with a time step of 2 fs and data capture every 0.5 ps. The RATTLE algorithm was used to restrain the high frequency bond stretch terms throughout all simulations. Prior equilibration was carried out for 400 ps with initial assignment of velocities according to the Boltzmann–Maxwell distribution and scaling by a single factor every 100 fs to maintain the temperature ±10 K. From the production phase, 89 structures were sampled, averaged and minimised as described. The MD simulation for the antibody–**5a$_{maj}$** complex contained a 1.3 ns production phase from which four average structures were minimised by sampling at four time points.

Acknowledgment

We thank Prof. W. H. Stimson (Department of Immunology, University of Strathclyde, Glasgow) for useful discussions on H11.

References and notes

1. (a) Schultz, P. G.; Yin, J.; Lerner, R. A. *Angew. Chem., Int. Ed.* **2002**, *41*, 4427–4437; (b) Blackburn, G. M.; Datta, A.; Denham, H.; Wentworth, P. *Adv. Phys. Org. Chem.* **1998**, *31*, 249–392; (c) Hilvert, D. *Annu. Rev. Biochem.* **2000**, *69*, 751–753; For a general review, see: (d) Keinan, E. In *Catalytic Antibodies*; Wiley-VCH: Weinheim, Germany, 2004.
2. (a) Maggiotti, V.; Resmini, M.; Gouverneur, V. *Angew. Chem., Int. Ed.* **2002**, *41*, 1012–1014; (b) Zhong, G. F.; Lerner, R. A.; Barbas, C. F. *Angew. Chem., Int. Ed.* **1999**, *38*, 3738–3741.
3. Chen, Y. W.; Reymond, J. L. *Synthesis-Stuttgart* **2001**, *6*, 934–936.
4. (a) Sinha, S. C.; Sun, J.; Wartmann, M.; Lerner, R. A. *Chembiochem* **2001**, *2*, 656–665; (b) List, B.; Shabat, D.; Zhong, G. F.; Turner, J. M.; Li, A.; Bui, T.; Anderson, J.; Lerner, R. A.; Barbas, C. F. *J. Am. Chem. Soc.* **1999**, *32*, 7283–7291; (c) Shabat, D.; List, B.; Lerner, R. A.; Barbas, C. F. *Tetrahedron Lett.* **1999**, *40*, 1437–1440.
5. Smithrud, D. B.; Benkovic, S. J.; Roberts, V.; Liu, J.; Neagu, I.; Phillips, B. W.; Smith, A. B.; Hirschmann, R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 1953–1958.
6. Hasserodt, J.; Janda, K. D.; Lerner, R. A. *J. Am. Chem. Soc.* **2000**, *122*, 40–45; (b) Hasserodt, J.; Janda, K. D.; Lerner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 5993–5998.
7. Hu, Y. J.; Ji, Y. Y.; Wu, Y. L.; Yang, B. H.; Yeh, M. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1601–1606.
8. Suckling, C. J.; Tedford, M. C.; Bence, L. M.; Irvine, J. I.; Stimson, W. H. *J. Chem. Soc., Perkin Trans. 1* **1993**, 1925–1929.
9. Pitt, A. R.; Stimson, W. H.; Suckling, C. J.; Marrero-Tellado, J. J.; Vazzana, C. *Isr. J. Chem.* **1996**, *36*, 171–175.
10. Khalaf, A. I.; Linaza, S.; Pitt, A. R.; Stimson, W. H.; Suckling, C. J. *Tetrahedron* **2000**, *56*, 489–495.
11. Brooks, L.; Suckling, C. J.; Stimson, W. H. *Biochem. Soc. Trans.* **1996**, *24*, 313S.
12. (a) Karlstrom, A.; Zhong, G.; Rader, C.; Larsen, N. A.; Heine, A.; Fuller, R.; List, B.; Tanaka, F.; Wilson, I. A.; Barbas, C. F., III; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3878–3883; (b) Gramatikova, S.; Mouratou, B.; Stetefeld, B.; Mehtac, P. K.; Christen, P. *J. Immunol. Methods* **2002**, 99–110; (c) Benedetti, F.; Berti, F.; Brady, K.; Colombatti, A.; Pauletto, A.; Pucillo, C.; Thomas, N. R. *Chembiochem* **2004**, *5*, 129–131.
13. Dall'Acqua, W.; Goldman, E. R.; Lin, W.; Teng, C.; Tsuchiya, D.; Li, H.; Ysern, X.; Braden, B. C.; Li, Y.; Smith-Gill, S. J.; Mariuzza, R. A. *Biochemistry* **1998**, *37*, 7981–7991.
14. Pokkuluri, P. R.; Raffen, R.; Dieckman, L.; Boogaard, C.; Stevens, F. J.; Schiffer, M. *Biophys. J.* **2002**, *82*, 391–398.
15. (a) Cross, S. S. J.; Brady, K.; Stevenson, J. D.; Sackin, J. R.; Kenward, N.; Dietel, A.; Thomas, N. R. *J. Immunol. Methods* **2002**, *269*, 173–195; (b) Nevanen, T. K.; Hellman, M. L.; Munck, N.; Wohlfahrt, G.; Koivula, A.; Soderlund, H. *Protein Eng.* **2003**, *16*, 1089–1097; (c) Guo, C. Z.; Wu, J. H.; Wang, Y. X.; Hu, Y. D.; Li, S.; Sun, M. J. *Acta Pharmacol. Sin.* **2003**, *24*, 460–466.
16. Janin, J.; Henrick, K.; Moult, J.; Eyck, L. T.; Sternberg, M. J. E.; Vajda, S.; Vakser, I.; Wodack, S. J. *Proteins* **2003**, *52*, 2–9.
17. Hex 4.2 User Manual available on-line at http://www.csd.abdn.ac.uk/hex/hex_manual.pdf.
18. Ritchie, D. W. *Proteins* **2003**, *52*, 98–106.
19. Ritchie, D. W.; Kemp, G. J. L. *Proteins* **2000**, *39*, 178–194.
20. Stryer, L. In *Biochemistry*, 4th ed.; WH Freeman and Company: New York, 1995; pp 368–372.
21. Cerius^2 version 4.8.1; Accelrys Inc. Cambridge UK.
22. Minasov, G.; Wang, X.; Shoichet, B. K. *J. Am. Chem. Soc.* **2002**, *124*, 5333–5340.
23. Mustafi, D.; Sosa-Peinado, A.; Makinen, M. W. *Biochemistry* **2001**, *40*, 2397–2409.
24. Tuganova, A.; Yoder, M. D.; Popov, K. M. *J. Biol. Chem.* **2001**, *276*, 17994–17999.
25. Bzymek, K. P.; Holz, R. C. *J. Biol. Chem.* **2004**, *279*, 31018–31025.
26. Debler, E. W.; Ito, S.; Seebeck, F. P.; Heine, A.; Hilvert, D.; Wilson, I. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 4984–4989.
27. Goncalves, O.; Dintinger, T.; Lebreton, J.; Blanchard, D.; Tellier, C. *Biochem. J.* **2000**, *346*, 691–698.
28. Golinelli-Pimpaneau, B.; Goncalves, O.; Dintinger, T.; Blanchard, D.; Knossow, M.; Tellier, C. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9892–9895.
29. Zheng, L.; Baumann, U.; Reymond, J.-L. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3387–3392.

30. (a) Heine, A.; Stura, E. A.; Yli-Kauhaluoma, J. T.; Gao, C.; Deng, Q.; Beno, B. R.; Houk, K. N.; Janda, K. D.; Wilson, I. A. *Science* **1998**, *279*, 1934–1940; (b) Hugot, M.; Bensel, N.; Vogel, M.; Reymond, M. T.; Stadler, B.; Reymond, J.-L.; Baumann, U. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9674–9678; (c) Zheng, L.; Goddard, J.-P.; Baumann, U.; Reymond, J.-L. *J. Mol. Biol.* **2004**, *341*, 807–814.
31. Goodman, J. M. In *Chemical Applications of Molecular Modelling*; The Royal Society of Chemistry: Cambridge, UK, 1998; pp 63–64.
32. Discovery Studio Modeler version 1.1; Accelrys Inc. Cambridge, UK., www.accelrys.com.
33. InsightII version 2000.3; Accelrys Inc. Cambridge, UK.
34. InsightII version 2000L; Accelrys Inc. Cambridge, UK.
35. MacBeath, G.; Hilvert, D. *Chem. Biol.* **1996**, *3*, 433–445.